

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2003, 01:10:18 ; Search time 18009 Seconds  
(without alignments)  
5843.509 Million cell updates/sec

Title: US-09-924-946-1  
Perfect score: 3616  
Sequence: 1 gattatgttgggggtcgc.....gttgctcattgactgct 3616

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_on.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3616	100.0	3616	6	AX460870	AX460870 Sequence
2	3529.8	97.6	3665	9	BC013153	BC013153 Homo sapi
3	2844.4	78.7	2976	6	AX323479	AX323479 Sequence
4	2665.6	73.7	2736	9	AY036093	AY036093 Sequence
5	2597.2	71.8	2603	6	AX377877	AX377877 Sequence
6	2443.6	67.6	3608	4	AF529202	AF529202 Bos tauru
7	2269.4	62.8	2271	9	AF338441	AF338441 Homo sapi
8	2265.4	62.6	2322	9	BC007522	BC007522 Homo sapi
9	2263.2	62.6	2268	6	AX323481	AX323481 Sequence
10	2258.2	62.5	2271	9	AF395336	AF395336 Homo sapi
11	2019.6	55.9	2059	9	AK025542	AK025542 Homo sapi
12	1730	47.8	2274	10	AF338440	AF338440 Mus muscu
13	1303.8	36.1	188192	9	AL139241	AL139241 Human DNA
14	1235.8	34.2	182222	2	AC123374	AC123374 Rattus no
15	853.8	23.6	3422	9	BC000594	BC000594 Homo sapi
16	849	23.5	3432	9	HSU89942	HSU89942 Human lysyl
17	767.4	21.2	2811	9	AF284815	AF284815 Homo sapi
18	765.8	21.2	2262	6	AR116983	AR116983 Sequence
19	765.8	21.2	2574	9	AF282619	AF282619 Homo sapi
20	765.8	21.2	2785	6	AX302669	AX302669 Sequence
21	765.8	21.2	2920	6	AR116982	AR116982 Sequence
22	765.8	21.2	3121	9	AF311313	AF311313 Homo sapi
23	738.4	20.4	3346	5	AF103901	AF103901 Perca fla
24	735.2	20.3	3429	5	AF143003	AF143003 Perca fla
25	707.6	19.6	2871	10	AF053368	AF053368 Mus muscu
26	706.2	19.5	3198	6	AX210644	AX210644 Sequence
27	706.2	19.5	3198	9	AF117949	AF117949 Homo sapi
28	706	19.5	2837	10	BC011298	BC011298 Mus muscu
29	450.2	12.5	460	6	AX408495	AX408495 Sequence
30	389.8	10.8	1689	9	AK098436	AK098436 Homo sapi
31	380.4	10.5	1394	9	BC033130	BC033130 Homo sapi
32	346	9.6	904	10	AF117951	AF117951 Mus muscu
33	322	8.9	327	6	AX041932	AX041932 Sequence
34	306.8	8.5	270516	2	AC096350	AC096350 Rattus no
35	248.6	6.9	227627	2	AC124552	AC124552 Mus muscu
36	239.6	6.6	513	6	AX210496	AX210496 Sequence
37	209.2	5.8	270516	2	AC096350	AC096350 Rattus no
38	199.8	5.5	1827	3	DME295626	DME295626 Drosophil
39	199.2	5.5	2476	6	AR142625	AR142625 Sequence
40	198.8	5.5	1917	4	AF421185	AF421185 Bos tauru
41	191.6	5.3	2328	9	HUMLYSOXLK	HUMLYSOXLK
42	191.6	5.3	2355	9	BC015090	BC015090 Homo sapi
43	190.6	5.3	1557	10	BC003973	BC003973 Mus muscu
44	190.6	5.3	2013	10	AF357006	AF357006 Mus muscu
45	184.6	5.1	187	6	AX041931	AX041931 Sequence

ALIGNMENTS

RESULT 1  
AX460870  
LOCUS AX460870 3616 bp DNA linear PAT 08-JUL-2002  
DEFINITION Sequence 1 from Patent WO0212470.  
ACCESSION AX460870  
VERSION AX460870.1 GI:21726122  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Evans, M.J., Scicchitano, M.S., Bapat, A.R., Beer, E., Bhat, R.A.,  
Ferris, E., Mastroeni, R., Zhang, J., and Karathanasis, S.K.  
TITLE A member of the lysyl oxidase gene family



```
Db 1981 CCCACTACGACCTCTCACTCTCAATGGCTCCAGGTGGCTGAGGGGCAACAAGGCCAGCT 2040
QY TCTGTCTGGAGACACAACACTGCCACACAGGACTGCAGCGGGCTACGCACTGTGCCAACT 2100
Db 2041 TCTGTCTGGAGACACAACACTGCCACACAGGACTGCAGCGGGCTACGCACTGTGCCAACT 2100
QY TTGAGAACAGGAGCTGACTGTAGGCTGTGGTGGGACACCTACCGGCATGACATTTGATTGCC 2160
Db 2101 TTGAGAACAGGAGCTGACTGTAGGCTGTGGTGGGACACCTACCGGCATGACATTTGATTGCC 2160
QY AGTGGTGGATATACAGATGTGGGCCCCGGGAATATATATCTTCCAGAGTGAATGTGAACC 2220
Db 2161 AGTGGTGGATATACAGATGTGGGCCCCGGGAATATATATCTTCCAGAGTGAATGTGAACC 2220
QY CCCACTATGAGTGGCAGAGTCAAGATTTCTCCAAACAATATGCTGCAGTGGCGCTGCAAGT 2280
Db 2221 CCCACTATGAGTGGCAGAGTCAAGATTTCTCCAAACAATATGCTGCAGTGGCGCTGCAAGT 2280
QY ATGATGGGCACCGGCTCTGGCTGCAACCTGCCACACAGGGAATTCATACCCAGCCAAATG 2340
Db 2281 ATGATGGGCACCGGCTCTGGCTGCAACCTGCCACACAGGGAATTCATACCCAGCCAAATG 2340
QY CAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACAACCTCATCTGAAGCTGTCACT 2400
Db 2341 CAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACAACCTCATCTGAAGCTGTCACT 2400
QY GCACACTCTAGCTGTGCTCGGATACACAGATACCTCAGCTTATTGGAGCCATGCCCTTC 2460
Db 2401 GCACACTCTAGCTGTGCTCGGATACACAGATACCTCAGCTTATTGGAGCCATGCCCTTC 2460
QY ACAGAGTCCCAACTCAGAGGAAAGGGCCAGTGCAGAGGGGACCAAGAACTGCTCAGG 2520
Db 2461 ACAGAGTCCCAACTCAGAGGAAAGGGCCAGTGCAGAGGGGACCAAGAACTGCTCAGG 2520
QY AAGCCTTTTGATGGCAGAGTACCAATCCAGATGGTATTGCTCCCTCAGSAGTGGCTGCG 2580
Db 2521 AAGCCTTTTGATGGCAGAGTACCAATCCAGATGGTATTGCTCCCTCAGSAGTGGCTGCG 2580
QY GCCTCCCTTAAGGGCTCTGGGCTATGGAATATGCTCCACAGGCTTTGCTCAGCTGAGC 2640
Db 2581 GCCTCCCTTAAGGGCTCTGGGCTATGGAATATGCTCCACAGGCTTTGCTCAGCTGAGC 2640
QY TCCTCTCTCTGAAGGAAACCCAGTCACTCCCTGAATCTTGCCACAGAGATCCGGGATTCAG 2700
Db 2641 TCCTCTCTCTGAAGGAAACCCAGTCACTCCCTGAATCTTGCCACAGAGATCCGGGATTCAG 2700
QY GAGCTCTCAGTTCTTTAGGGATGGAATGATGCTCCACAGGCTTTGCTCAGCTGAGC 2760
Db 2701 GAGCTCTCAGTTCTTTAGGGATGGAATGATGCTCCACAGGCTTTGCTCAGCTGAGC 2760
QY AAATGCTTTGGAGGAGTATAGGACAGAGGACCAAAATACACAGCAGGTAGTGTGCTCT 2820
Db 2761 AAATGCTTTGGAGGAGTATAGGACAGAGGACCAAAATACACAGCAGGTAGTGTGCTCT 2820
QY CTGCTAGGAGCTCAAGGACACAACTTGATCAAAATCAAACTGGCAGAGAGCTGGT 2880
Db 2821 CTGCTAGGAGCTCAAGGACACAACTTGATCAAAATCAAACTGGCAGAGAGCTGGT 2880
QY GGATCCAACTCTTCTTCACTCTCTGTTATTAGAACTCACTCTCAGCTCTGTCTCTTT 2940
Db 2881 GGATCCAACTCTTCTTCACTCTCTGTTATTAGAACTCACTCTCAGCTCTGTCTCTTT 2940
QY AGTGTCTTTACCTTTATCTTACACACACATGGGTGTTTCTATTATCTTGGAGACACAG 3000
Db 2941 AGTGTCTTTACCTTTATCTTACACACACATGGGTGTTTCTATTATCTTGGAGACACAG 3000
QY ACCTGGGCACTCCCTTATGCTGTGAGTGGGCAACCAACAGTACGAGTGTGTGAGA 3060
Db 3001 ACCTGGGCACTCCCTTATGCTGTGAGTGGGCAACCAACAGTACGAGTGTGTGAGA 3060
QY AGGGGCAAGTTTCAGAAATGGCCAGATAGGCGCTTCTTACAGAGCAGCAGATAGGC 3120
Db 3061 AGGGGCAAGTTTCAGAAATGGCCAGATAGGCGCTTCTTACAGAGCAGCAGATAGGC 3120
```

```
Db 3061 AGGGGCAAGTTTCAGAAATGGCCAGATAGGCGCTTCTTACAGAGCAGCAGATAGGC 3120
QY CAAGCAGAAAGACTGTCTGAGGTAAACAGGAGACCCAGCCCTCTGAGGGCTCTGCCAAGG 3180
Db 3121 CAAGCAGAAAGACTGTCTGAGGTAAACAGGAGACCCAGCCCTCTGAGGGCTCTGCCAAGG 3180
QY AAATAATATGGACCATTTACTCTGGCAGCAGTGTCTCTCTCAGGATCAACACGATC 3240
Db 3181 AAATAATATGGACCATTTACTCTGGCAGCAGTGTCTCTCTCAGGATCAACACGATC 3240
QY TCAGGATTTGCTTAAACTTCAAGTCTCAACCAAGTGTCTGAAGTGAACCTTTGCAATTGAAT 3300
Db 3241 TCAGGATTTGCTTAAACTTCAAGTCTCAACCAAGTGTCTGAAGTGAACCTTTGCAATTGAAT 3300
QY AAATTTTGGCCATGGAAAGAACATCAACCAAGCCACTCATCTCTACAGAGATAAGAAAAAC 3360
Db 3301 AAATTTTGGCCATGGAAAGAACATCAACCAAGCCACTCATCTCTACAGAGATAAGAAAAAC 3360
QY AACTTTTGGCAGACAGACAGACAGAACGCTGGAGAAATCAGAAGGGGGAACAGTCAGTT 3420
Db 3361 AACTTTTGGCAGACAGACAGACAGAACGCTGGAGAAATCAGAAGGGGGAACAGTCAGTT 3420
QY TAGTTAAGSAGTGAACCTGGGAAAGGCCACCATTCCTGCTTGTGATGGGCTCTGATTTGCT 3480
Db 3421 TAGTTAAGSAGTGAACCTGGGAAAGGCCACCATTCCTGCTTGTGATGGGCTCTGATTTGCT 3480
QY CTTGCTCAAGTGGAAATAAAACCCCATGCTTTTGGACATGATTTCTTGATCTTTTCTCCA 3540
Db 3481 CTTGCTCAAGTGGAAATAAAACCCCATGCTTTTGGACATGATTTCTTGATCTTTTCTCCA 3540
QY CTGAGACACACTTAAAGTGATGATCTTACAGAGCTGACACCCCTAATGCCAATAAAGTTG 3600
Db 3541 CTGAGACACACTTAAAGTGATGATCTTACAGAGCTGACACCCCTAATGCCAATAAAGTTG 3600
QY CTCATTATGAGTCTCT 3616
Db 3601 CTCATTATGAGTCTCT 3616
```

RESULT 2  
BC013153  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BC013153  
Homo sapiens, clone MGC:17373 IMAGE:3872314, mRNA, complete cds.  
BC013153  
BC013153.1 GI:15341930  
MGC.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 3665)  
Strausberg,R.  
Direct Submission  
Submitted (27-AUG-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalob@bcm.tmc.edu](mailto:villalob@bcm.tmc.edu)  
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 13 Row: 1 Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: GenomeScan gene  
 prediction.

## FEATURES

source Location/Qualifiers

1..3665  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:17373 IMAGE:3872314"  
 /tissue="Eye, retinoblastoma"  
 /clone\_lib="NIH\_MGC\_67"  
 /lab\_host="DH10B"  
 /notes="vector: pCMV-Sport6"  
 152..2422  
 /codon\_start=1  
 /product="Unknown (protein for MGC:17373)"  
 /protein\_id="AAH13153.1"  
 /db\_xref="GI:15341931"

## CDS

/translation="MARSPATLFLFLLLLGPPSPFQSIGTKLRILVGPESKPBEG  
 RLEVLHOGQGTVCDDNPAIQEATVACQLGFEEALTAHSAKYGGEGPIHLNDYQC  
 VGTESLDQCGSNMGVSDSHSDVGIHPRHRGYSLETYSNALGQRRLEEV  
 LKPIASAKOHSVTEGAVEVKEGHMROVDQGWMTNNSRVVCCMLGPPSEVPVDSH  
 YVRKMDLKWDRPKSLTKNSFWTHQVTLCTEPMANCOVQVAFARGKLRPAC  
 PGMHIAVSCVAGPHRPPTKPKQKGSWABEPRVLRSGAOVGRVRLVLRNROGT  
 VCDHRNLIASVVCRLQFGSAREALFGRLLGSLGPIHLSEVRGRIYTLSDCPA  
 LEGSQNCQHENAAVRNPNMGPQVRLAGRIPEGLLEVEVNVNGVPRWGSVC  
 SENMGTEAMVACRLQLGFALHAYKETWFSNGTTPRAEVNMSVRCSTELALQCCQ  
 RGHVHCSSGGRGLAGVSCMDSAPDLVMAQLVOETAYLEDRLPSOLYCAHEENCL  
 KADHMDWPGYRLLRFPSTQIYNLGRDPRKTRDQSWHQCRRHYSIEVFTHYD  
 LLTLNKSVAEGHKAFCLEDTNPTGLORRYACANFGQVTVGWDYRHDIDCW  
 VDLTDPGNYIFQVINYPHYVEAESDFSNMLQCRKYDGRHVRVHLNCHTGNISYPAN  
 AELSLEBEQRLRNLI"

BASE COUNT 824 a 1021 c 1085 g 735 t  
 ORIGIN

Query Match 97.6%; Score 3529.8; DB 9; Length 3665;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 3540; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
 QY 60 CAAGCCCTCTTCTTAGCCTGACAGATATCTTGGCTCAGCTCTCTTGAAGTCAAC 119  
 DB 92 CAGCCCCGAGCTCCGCGCTCCATCTGGTATCTTGGCTCAGCTCTCTTGAAGTCAAC 151  
 QY 120 ATGCGTGTGTCACACAGCCACCTCTTTCTGTCTCTGCTGCTAGGCCAGCCCT 179  
 DB 152 ATGCGAGGTCTCCACAGCCACCTCTTTCTGTCTCTGCTGCTAGGCCAGCCCT 211  
 QY 180 CCCAGAGGCCACAGTCACTGGGACCACTAAGCTCCGGCTGGTGGGCCACAGAGCAAG 239  
 DB 212 CCCAGAGGCCACAGTCACTGGGACCACTAAGCTCCGGCTGGTGGGCCACAGAGCAAG 271  
 QY 240 CCAGAGAGGCCCGCTGAGGTGTGTACACAGGGCCAGTGGGGACCGCTGTGTATGAC 299  
 DB 272 CCAGAGAGGCCCGCTGAGGTGTGTACACAGGGCCAGTGGGGACCGCTGTGTATGAC 331  
 QY 300 AACTTTGCTATCCAGAGGCCACAGTGGCTTCCGGCCAGTGGGCTTCCAGCTGCTTG 359  
 DB 332 AACTTTGCTATCCAGAGGCCACAGTGGCTTCCGGCCAGTGGGCTTCCAGCTGCTTG 391  
 QY 360 ACCTGGGCCACAGTGGCCAGTAAGCCAGAGGGAGGACCCATCTGTGTGTGACATGTG 419  
 DB 392 ACCTGGGCCACAGTGGCCAGTAAGCCAGAGGGAGGACCCATCTGTGTGACATGTG 451  
 QY 420 CGCTGTGTGGCCACAGAGCTCTTGGACCACTGGGGTCTTAATGGCTGGGGAGTCACT 479  
 DB 452 CAGTGTGTGGCCACAGAGCTCTTGGACCACTGGGGTCTTAATGGCTGGGGAGTCACT 511  
 QY 480 GACTGTGAGTCACTCAGAGAGCTAGGGGTGATATGACCCCGCCGCTCTGTGGTAC 539  
 DB 512 GACTGTGAGTCACTCAGAGAGCTAGGGGTGATATGACCCCGCCGCTCTGTGGTAC 571  
 QY 540 CTTTCTGAAACTGTCTCAATGCCCTTGGGCCCGGCGGCTGGAGGAGGTGCGG 599

DB 572 CTTTCTGAAACTGTCTCAATGCCCTTGGGCCCGGCGGCTGGAGAGGTGCGG 631  
 QY 600 CTCAGAGCCATCTTGTCCAGTGCAGAGCAGCAGTACGACCTGACCCAGGAGGAGCCGTGGAG 659  
 DB 632 CTCAGAGCCATCTTGTCCAGTGCAGAGCAGCAGTACGACCTGACCCAGGAGGAGCCGTGGAG 691  
 QY 660 GTGAAGTATGAGGGCCATCTGGGGCAGGTGTGTGACAGGCTGTGGACCATCAACAAACAGC 719  
 DB 692 GTGAAGTATGAGGGCCATCTGGGGCAGGTGTGTGACAGGCTGTGGACCATCAACAAACAGC 751  
 QY 720 AGGTGTGTGCGGGATCTGGGCTTCCCGAGGAGGTGCTGTGCGACAGCAGCTACTTAC 779  
 DB 752 AGGTGTGTGCGGGATCTGGGCTTCCCGAGGAGGTGCTGTGCGACAGCAGCTACTTAC 811  
 QY 780 AGGAAGTCTGGGATCTGAAGATGAGGGACCTTAAGTCTTGGCTGAAGAGCCTGACGAAT 839  
 DB 812 AGGAAGTCTGGGATCTGAAGATGAGGGACCTTAAGTCTTGGCTGAAGAGCCTGACGAAT 871  
 QY 840 AAGAACTCTTCTGATCCACAGGTCACTGCTGTGGGGCAGAGCCCGCAGTGGCCAAAC 899  
 DB 872 AAGAACTCTTCTGATCCACAGGTCACTGCTGTGGGGCAGAGCCCGCAGTGGCCAAAC 931  
 QY 900 TGCCAGGTGACGTGCTCCAGCCCGGGCAAGCTGCGGK CAGCTGCCCCAGGTGGCATG 959  
 DB 932 TGCCAGGTGACGTGCTCCAGCCCGGGCAAGCTGCGGK CAGCTGCCCCAGGTGGCATG 991  
 QY 960 CATGCTGTGTGAGTGTGTGGAGGCTCTCACTTCCGCK CACCAGAGCAAGAGCCACAA 1019  
 DB 992 CATGCTGTGTGAGTGTGTGGAGGCTCTCACTTCCGCK CACCAGAGCAAGAGCCACAA 1051  
 QY 1020 CGCAAGGTCTCTGGGACAGGAGCCAGGCTGCGCTGK GCTCCGGGGCCAGGTGGGK 1079  
 DB 1052 CGCAAGGTCTCTGGGACAGGAGCCAGGCTGCGCTGK GCTCCGGGGCCAGGTGGGK 1111  
 QY 1080 GAGGCCCGGGTGGAAAGTGTCTAAGAACCCGCAAGTGGGGK CGGTCTGTGACACAGGTGG 1139  
 DB 1112 GAGGCCCGGGTGGAAAGTGTCTAAGAACCCGCAAGTGGGGK CGGTCTGTGACACAGGTGG 1171  
 QY 1140 AACCTCATCTCTGCCAGTGTGTGTGCTGCTGAGCTGGGCTTGGCTCTGCTGGGAGGCC 1199  
 DB 1172 AACCTCATCTCTGCCAGTGTGTGTGCTGCTGAGCTGGGCTTGGCTCTGCTGGGAGGCC 1231  
 QY 1200 CTCTTTGGGGCCCGGCTGGGCGCAAGGGCTAGGGCCCATCACTGAGTGAAGTGGCTGCTG 1259  
 DB 1232 CTCTTTGGGGCCCGGCTGGGCGCAAGGGCTAGGGCCCATCACTGAGTGAAGTGGCTGCTG 1291  
 QY 1260 AGGGGATATGAGGGAGCCCTCAGCGACTGCGCTGCGCTGCAAGGGTCCAGAAATGGTTGC 1319  
 DB 1292 AGGGGATATGAGGGAGCCCTCAGCGACTGCGCTGCGCTGCAAGGGTCCAGAAATGGTTGC 1351  
 QY 1320 CAACATGAAATGCTGCTGTGCTGCAATGTCCCTAACATGGGCTTTTTCAGAAATCAG 1379  
 DB 1352 CAACATGAAATGCTGCTGTGCTGCAATGTCCCTAACATGGGCTTTTTCAGAAATCAG 1411  
 QY 1380 GTGCGCTTTGGCTGGTGGGCTATCCCTGAGGAGGGGCTA TGGAGGTGCAAGTGGAGGTG 1439  
 DB 1412 GTGCGCTTTGGCTGGTGGGCTATCCCTGAGGAGGGGCTA TGGAGGTGCAAGTGGAGGTG 1471  
 QY 1440 AACGGGTCCACAGCTGGGGAGCGTGTGCAGTGAAGAAC TGGGGCTCACCGAGGCCATG 1499  
 DB 1472 AACGGGTCCACAGCTGGGGAGCGTGTGCAGTGAAGAAC TGGGGCTCACCGAGGCCATG 1531  
 QY 1500 GTGGCTGCGGACAGCTCGGCTCTGGGTCTTGGCCATCATTCCTAACAGGAAACCTGGTTTC 1559  
 DB 1532 GTGGCTGCGGACAGCTCGGCTCTGGGTCTTGGCCATCATTCCTAACAGGAAACCTGGTTTC 1591  
 QY 1560 TGGTGGGGACCGCAAGGGCCAGAGGCTGATGAGTGGGTGGCTGCTGCTCAGGCACA 1619  
 DB 1592 TGGTGGGGACCGCAAGGGCCAGAGGCTGATGAGTGGGTGGCTGCTGCTCAGGCACA 1651  
 QY 1620 GAGCTGGGCTCAGCAGTGCAGAGGACCGGGCGGTG TACTGCTCCACCGGTGGCGGG 1679

Db 1652 GAGCTGGCCCTGCAGAGTGCACAGAGGACGGCCGGGTGCACTGCTCCCCAGGTGGCGG 1711  
Qy CGCTTCTGGCTGGAGTCTCCATGATGAGACAGGACCTGGTATGAAAGCCCGCAG 1739  
Db 1712 CGCTTCTGGCTGGAGTCTCCATGATGAGACAGGACCTGGTATGAAAGCCCGCAG 1771  
Qy 1740 CTAGTGCAGGACAGGCTTCTTGGAGGACCGCCGCTAGCCAGCTGTATTTGGCCAC 1799  
Db 1772 CTAGTGCAGGACAGGCTTCTTGGAGGACCGCCGCTAGCCAGCTGTATTTGGCCAC 1831  
Qy 1800 GAGGAACTGCTCTCCAGTCTGGGATCAGATGAGTGGCCCTACGGATACGGCCGC 1859  
Db 1832 GAGGAACTGCTCTCCAGTCTGGGATCAGATGAGTGGCCCTACGGATACGGCCGC 1891  
Qy 1860 CTATTCGCTTCTCCACAGATCTCAATCTGGCCGGAGTGTCTTCCAGAGACT 1919  
Db 1892 CTATTCGCTTCTCCACAGATCTCAATCTGGCCGGAGTGTCTTCCAGAGACT 1951  
Qy 1920 GGACCGATAGTGGGTTTGGACCAAGTGCACAGGCAATACCAAGATTTGAGGTCTTC 1979  
Db 1952 GGACCGATAGTGGGTTTGGACCAAGTGCACAGGCAATACCAAGATTTGAGGTCTTC 2011  
Qy 1980 ACCCACTAGGACTCTCTCACTCTCAATGGCTCCAAAGTGGCTGAGGGGSCAACAGGCCAGC 2039  
Db 2012 ACCCACTAGGACTCTCTCACTCTCAATGGCTCCAAAGTGGCTGAGGGGSCAACAGGCCAGC 2071  
Qy 2040 TTCTGTGTGGAGACAAACTGCCCAAGAGCTGAGGGGCTGAGGGGSCAACAGGCCAGC 2099  
Db 2072 TTCTGTGTGGAGACAAACTGCCCAAGAGCTGAGGGGCTGAGGGGSCAACAGGCCAGC 2131  
Qy 2100 TTTGAGAACAGGAGTGTAGTGTAGTGTGGACACCTACCGGCATGACATTTGATTC 2159  
Db 2132 TTTGAGAACAGGAGTGTAGTGTAGTGTGGACACCTACCGGCATGACATTTGATTC 2191  
Qy 2160 CAGTGGGTGATACACAGATGGGCGCGGAAATATATCTTCCAGGTGATTTGAAAC 2219  
Db 2192 CAGTGGGTGATACACAGATGGGCGCGGAAATATATCTTCCAGGTGATTTGAAAC 2251  
Qy 2220 CCCCACTATGAAGTGGCAGAGTGCAGATTTCTCCAAACAAATATGCTGCGAGTGGCGTGAAG 2279  
Db 2252 CCCCACTATGAAGTGGCAGAGTGCAGATTTCTCCAAACAAATATGCTGCGAGTGGCGTGAAG 2311  
Qy 2280 TATGATGGGACCGGGTCTGGTGTGCAACTGCGCACACAGGAAATTCATCCAGCCCAAT 2339  
Db 2312 TATGATGGGACCGGGTCTGGTGTGCAACTGCGCACACAGGAAATTCATCCAGCCCAAT 2371  
Qy 2340 GCAGAACTCTCCGTGAGCAGGAAACAGCGTCTCAGGAAACAACTCATCTGAAGCTGTAC 2399  
Db 2372 GCAGAACTCTCCGTGAGCAGGAAACAGCGTCTCAGGAAACAACTCATCTGAAGCTGTAC 2431  
Qy 2400 TGCACACTCTAGCTGCTCCGATACACAGATACCTCAGCTTATTTGGAGCCATGCCCTT 2459  
Db 2432 TGCACACTCTAGCTGCTCCGATACACAGATACCTCAGCTTATTTGGAGCCATGCCCTT 2491  
Qy 2460 CACAGAGTCCCAACTCAGAGGAAAGGCGCAGTGCAGGAGGACCAAGAACCTGCTCAG 2519  
Db 2492 CACAGAGTCCCAACTCAGAGGAAAGGCGCAGTGCAGGAGGACCAAGAACCTGCTCAG 2551  
Qy 2520 GAAGCCTTTTGTATGGCAAGATCAACCAATCCAGATGGTATTTGCTCCCTCAGGATGGCTCTG 2579  
Db 2552 GAAGCCTTTTGTATGGCAAGATCAACCAATCCAGATGGTATTTGCTCCCTCAGGATGGCTCTG 2611  
Qy 2580 GGCCTGCCCTTAAGGCGCTGTGGCTATGAAATATGCTCCAGGCTTTGTACGCTGAG 2639  
Db 2612 GGCCTGCCCTTAAGGCGCTGTGGCTATGAAATATGCTCCAGGCTTTGTACGCTGAG 2671  
Qy 2640 CTCCTCTTCTGAAGGAAACCCAGTCACTCCTGATCTTGGCCACAGATCCGGGATTC 2699  
Db 2672 CTCCTCTTCTGAAGGAAACCCAGTCACTCCTGATCTTGGCCACAGATCCGGGATTC 2731  
Qy 2700 GGAGCTCTCAGTTCTTATGGGATGGATGATGGCCCGAGTCCCGCATCTAAGTGGTCTTTG 2759  
Db 2732 GGAGCTCTCAGTTCTTATGGGATGGATGATGGCCCGAGTCCCGCATCTAAGTGGTCTTTG 2791

Qy 2760 CAAATGTCTTTGGAGGAGTATAGACAGAGGACCAAAATACACAGAGCTAGTGTAGCTC 2819  
Db 2792 CAAATGTCTTTGGAGGAGTATAGACAGAGGACCAAAATACACAGAGCTAGTGTAGCTC 2851  
Qy 2820 TCTGCTAGGAGCTCAAAAGCAACAACTTGTATCAAAATCAAACTGGCAGAGAGCTGG 2879  
Db 2852 TCTGCTAGGAGCTCAAAAGCAACAACTTGTATCAAAATCAAACTGGCAGAGAGCTGG 2911  
Qy 2880 TGGATCCAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2939  
Db 2912 TGGATCCAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2971  
Qy 2940 TATGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2999  
Db 2972 TATGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3031  
Qy 3000 GACCTCGGGCATCCCTTATTTGGCTGTATGGGCAACCAACAGTTTACGGAGTCTTGG 3059  
Db 3032 GACCTCGGGCATCCCTTATTTGGCTGTATGGGCAACCAACAGTTTACGGAGTCTTGG 3091  
Qy 3060 AAGGGCAAGTTTACAGAAATGGCCAGATAGGGCTTCTTACAGAGCAGCAAGAGTAGG 3119  
Db 3092 AAGGGCAAGTTTACAGAAATGGCCAGATAGGGCTTCTTACAGAGCAGCAAGAGTAGG 3151  
Qy 3120 CCAAGCAGAAAGACTGCTGAGGTAAACAGGACCCAGGCCCTGTACGGGCTTGCAG 3179  
Db 3152 CCAAGCAGAAAGACTGCTGAGGTAAACAGGACCCAGGCCCTGTACGGGCTTGCAG 3211  
Qy 3180 GAAATATATGGACATTTACCTGGCAGGAGTGTCTCTCTCAGGATCACACGAT 3239  
Db 3212 GAAATATATGGACATTTACCTGGCAGGAGTGTCTCTCTCAGGATCACACGAT 3271  
Qy 3240 CTGAGATTTGGTCTTAACTTCAAGTCTCAACCAAGTGTCTGAAGTGAATTTGCAATTGAA 3299  
Db 3272 CTGAGATTTGGTCTTAACTTCAAGTCTCAACCAAGTGTCTGAAGTGAATTTGCAATTGAA 3331  
Qy 3300 TAAATTTTCCATGGAAAGCAATCAACCAAGCCACTCATCTTACAGAGATTAAGAAA 3359  
Db 3332 TAAATTTTCCATGGAAAGCAATCAACCAAGCCACTCATCTTACAGAGATTAAGAAA 3391  
Qy 3360 CAAGTTTGGCAGAGCAGAGACAGAGACCGTGGAGAAATCAGAGGGGGAACAGTCAGT 3419  
Db 3392 CAAGTTTGGCAGAGCAGAGACAGAGACCGTGGAGAAATCAGAGGGGGAACAGTCAGT 3451  
Qy 3420 TTAGTTAAGGATGGAACCTGGGAAAGGCCCACTTCTGCTTGTAGTGGGCTCTGATTGC 3479  
Db 3452 TTAGTTAAGGATGGAACCTGGGAAAGGCCCACTTCTGCTTGTAGTGGGCTCTGATTGC 3511  
Qy 3480 TCTTGTCTCAAGTGAATTAACCAACCCCATGCTTCTTGTGACATGATCTTGTATCTTTCTCC 3539  
Db 3512 TCTTGTCTCAAGTGAATTAACCAACCCCATGCTTCTTGTGACATGATCTTGTATCTTTCTCC 3571  
Qy 3540 ACTGAGACACACTTAAAGTGTATGCTTACAGGACTGACACCCCTAATGCCAATAAAGTT 3599  
Db 3572 ACTGAGACACACTTAAAGTGTATGCTTACAGGACTGACACCCCTAATGCCAATAAAGTT 3631  
Qy 3600 GCTCATTATGAGTGTCT 3616  
Db 3632 GCTCATTATGAGTGTCT 3648

RESULT 3  
AX323479  
LOCUS AX323479 2976 bp DNA linear PAT 07-JAN-2002  
DEFINITION Sequence 1 from Patent WO0192495.  
ACCESSION AX323479  
VERSION AX323479.1 GI:18094234  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.





```
QY 1740 CTAGTGCAGGAGACGGCCCTACTTTGAGGACCGCCCGCTCAGCCAGCTGTATTGTGCCAC 1799
Db 1715 CTAGTGCAGGAGACGGCCCTACTTTGAGGACCGCCCGCTCAGCCAGCTGTATTGTGCCAC 1774
QY 1800 GAGGAGAACTGCTCTCMAAGTCTCGGGATCAGATGAGCTGGCCCTACGGATACCGCGC 1859
Db 1775 GAGGAGAACTGCTCTCMAAGTCTCGGGATCAGATGAGCTGGCCCTACGGATACCGCGC 1834
QY 1860 CTATTGCGCTCTCCACACAGATCTACATCTGGCGGAGCTGACTTTTCGTCCTCAAGACT 1919
Db 1835 CTATTGCGCTCTCCACACAGATCTACATCTGGCGGAGCTGACTTTTCGTCCTCAAGACT 1894
QY 1920 GGACGCGATAGTGGTGGTGGACACAGTGGCCACAGGCAATACACAGCAATTTGAGGTCTTC 1979
Db 1895 GGACGCGATAGTGGTGGTGGACACAGTGGCCACAGGCAATTTGAGGTCTTC 1954
QY 1980 ACCCACTACGACCTCTCACTCTCAATGGCTCCAAAGTGGCTGAGGGGCAACAGGCCAGC 2039
Db 1955 ACCCACTACGACCTCTCACTCTCAATGGCTCCAAAGTGGCTGAGGGGCAACAGGCCAGC 2014
QY 2040 TTCTGCTGGAGGACACAACTGCCCCACAGGACTGCGAGCGGCGCTTACGATGTGCCAAC 2099
Db 2015 TTCTGCTGGAGGACACAACTGCCCCACAGGACTGCGAGCGGCGCTTACGATGTGCCAAC 2074
QY 2100 TTTGGAGAACAGGGAGTGAAGTCTAGGCTGCTGGACACCTTACCGGCATGACATTTGATTC 2159
Db 2075 TTTGGAGAACAGGGAGTGAAGTCTAGGCTGCTGGACACCTTACCGGCATGACATTTGATTC 2134
QY 2160 CAGTGGTGGATATCACAGATGTGGGCGGGAATTTATATCTTCAGGCTGATTTGTAAC 2219
Db 2135 CAGTGGTGGATATCACAGATGTGGGCGGGAATTTATATCTTCAGGCTGATTTGTAAC 2194
QY 2220 CCCCACTATGAAGTGGCAGAGTCAAGTTTCTCCAAACAAATATGCTGCAAGTGGCGTGAAG 2279
Db 2195 CCCCACTATGAAGTGGCAGAGTCAAGTTTCTCCAAACAAATATGCTGCAAGTGGCGTGAAG 2254
QY 2280 TATGATGGCCACCGGTCTGGTGTGCAAACTGCGACACAGGGAATTCATACCCAGCCAAAT 2339
Db 2255 TATGATGGCCACCGGTCTGGTGTGCAAACTGCGACACAGGGAATTCATACCCAGCCAAAT 2314
QY 2340 GCAGAACTCTCCTGGAGCAGGACAGCGTCTCAGGAAACCTCATCTGAAGCTGTAC 2399
Db 2315 GCAGAACTCTCCTGGAGCAGGACAGCGTCTCAGGAAACCTCATCTGAAGCTGTAC 2374
QY 2400 TGCACACTCTAGCTGCTCGGATACACAGATACCTCAGCTTATTCGAGCATGCGCTT 2459
Db 2375 TGCACACTCTAGCTGCTCGGATACACAGATACCTCAGCTTATTCGAGCATGCGCTT 2434
QY 2460 CACAGAGTCCCAACTCAGAGGAAAGGGCCAGTGCCAAAGGGGACCAAGAACCTGCTCAG 2519
Db 2435 CACAGAGTCCCAACTCAGAGGAAAGGGCCAGTGCCAAAGGGGACCAAGAACCTGCTCAG 2494
QY 2520 GAAGCCTTTTATGGAAGATCAACCAATCCAGATGATTTGCTCCCTCAGGATGGCTCTG 2579
Db 2495 GAAGCCTTTTATGGAAGATCAACCAATCCAGATGATTTGCTCCCTCAGGATGGCTCTG 2554
QY 2580 GGCCTGCCCTTAAGGCGCTGCGCTTATGGAATATGCTTCCAGGCTTTGCTAGCTGAG 2639
Db 2555 GGCCTGCCCTTAAGGCGCTGCGCTTATGGAATATGCTTCCAGGCTTTGCTAGCTGAG 2614
QY 2640 CTCCTCTTCTGAAGAAACCCAGTCACTCCTGATCTTGGCCACAGAGATCCGGGATTC 2699
Db 2615 CTCCTCTTCTGAAGAAACCCAGTCACTCCTGATCTTGGCCACAGAGATCCGGGATTC 2674
QY 2700 GGAGCTCTCAGTTCTTAGGGATGGACTATGGCCAGTCCCGCATCTAAGTGGTCTTTG 2759
Db 2675 GGAGCTCTCAGTTCTTAGGGATGGACTATGGCCAGTCCCGCATCTAAGTGGTCTTTG 2734
QY 2760 CAAATCTCTTGGAGGAGTATGACACAGAGACCAAAATACACAGCAGTGTGTTAGCTC 2819
Db 2735 CAAATCTCTTGGAGGAGTATGACACAGAGACCAAAATACACAGCAGTGTGTTAGCTC 2794
```

```
QY 2820 TCTGTAGAGCTCAAGCAACCAACTTGTATCAAAATCACACTGGCAGAGAACTGG 2879
Db 2795 TCTGTAGAGCTCAAGCAACCAACTTGTATCAAAATCACACTGGCAGAGAACTGG 2854
QY 2880 TGGATCAATCCCTTCTTCACTCTGTTGTTATTAGAACTCACCTCTCACCTCTGTTCTT 2939
Db 2855 GTGGATCAATCCCTTCTTCACTCTGTTGTTATTAGAACTCAACCTTCACTCTGTTCTT 2914
QY 2940 TAGTGTCTTACCTTATCTTACCAACACACATGGTGTGTTTATTATTCTTGAAGCACA 2999
Db 2915 TAGGCGCTTACCTTATCTTATTTWACACAMAAATNGNGGTTTTATTATCTTGTGAAGCACA 2974
QY 3000 GA 3001
Db 2975 AA 2976

LOCUS AY036093 2736 bp mRNA linear PRI 15-DEC-2001
DEFINITION Homo sapiens lysyl oxidase-like 4 mRNA, complete cds.
ACCESSION AY036093
VERSION AY036093.1 GI:17861371
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2736)
AUTHORS Maki, J.M., Tikkanen, H. and Kivirikko, K.I.
TITLE Cloning and characterization of a fifth human lysyl oxidase
isoenzyme: the third member of the lysyl oxidase-related subfamily
with four scavenger receptor cysteine-rich domains
JOURNAL Matrix Biol. 20 (7), 493-496 (2001)
MEDLINE 21550107
PubMed 11691589
REFERENCE 2 (bases 1 to 2736)
AUTHORS Maki, J.M., Tikkanen, H. and Kivirikko, K.I.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-2001) Department of Medical Biochemistry,
University of Oulu, P.O. Box 5000, Oulu 90014, Finland

FEATURES
source
1..2736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
7..2277
/codon_start=1
/product="lysyl oxidase-like 4"
/protein_id="AAK64186.1"
/db_xref="GI:17861372"
/translation="MAWSPATLFLFLLLLGQPPSRPQSLGTTKLRLVGPESKPEEG
RLVHLHQMGVTCDDNFALQEAATVACQKQFALTAHSAKYGQGGEGFIDLNVRC
VTGESSLDQGSNGVSDSHEDVGHCHPRHRGYSLETNSNALGQGRLEEVH
LKPILASAKQHSFVTEGAVEYGVHWRQVQDMNNSRVVCGMLGPPSEVPDVS
YRKWDLKMRDPKSLTKNSFWHTQVCLCTEPHMANCOVAPARGKLRLPAC
PGMHAVVSCVAGPHERPKTPQKQSWAEEPRVLSGAQVGEGRVLMNRQGT
VCDHRNWLISASVVRQGFSGAREALFGALGGQGLPIHLSEVRCRVERTLSDCPA
LEGSQGCQHENDAAVRCNPMQFQNGRIPBEGLELVGVNMGVPRWQSVY
SENGLTEAMVACQKQGLFAIHAYKETFWSGTPRAQEVNMGVRCSGTELALQCCQ
RHCPVHCSCGGRFLAGVSCMDAPDLVMAQLVQETAYLEDRPLSOLYCAHECNLS
KSDHMDNPGYRRLRFSTQIYNLGRDTPKTRGDSKWHQCHRHHSILEVTHVD
LITLNGSKVAEGHAKSFCELDNCTGLQRYACANFGQGVGVGWDYRHDIDCW
VDITDVGPNGYIFQVIVNPHYEVAESDFSNMLQCRCKYDGHVRVLIHNCHTNSYPAN
AELSLEQRLRNLI"

BASE COUNT 565 a 788 c 862 g 521 t
ORIGIN
Query Match 73.7%; Score 2665.6; DB 9; Length 2736;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2692; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 114 GTCACATAGGCGTGGTCCCAACAGCCACCTCTTCTTCTGCTGCTAGGCCAG 173
```

Db 1 GTCCACATGCGTGGTCCCCACAGCCACCCCTCTTTCTGTTCTGCTGCTAGGCCAG 60  
Qy 174 CCCCCTCCACAGCCACAGTCACTAGGCGCCACCTAAGCTCGGCTGGTGGGCCACAG 233  
Db 61 CCCCCTCCACAGCCACAGTCACTAGGCGCCACCTAAGCTCGGCTGGTGGGCCACAG 120  
Qy 234 AGCAAGCCACAGAGGGCGCTGGAGGTGTGACACAGGGCCAGTGGGGCACCGTGTGT 293  
Db 121 AGCAAGCCACAGAGGGCGCTGGAGGTGTGACACAGGGCCAGTGGGGCACCGTGTGT 180  
Qy 294 GATGACAACTTTGTATCCAGAGGCCACAGTGGCTTGCAGCAGCTGGGCTTTCGAAGCT 353  
Db 181 GATGACAACTTTGTATCCAGAGGCCACAGTGGCTTGCAGCAGCTGGGCTTTCGAAGCT 240  
Qy 354 GCCTTGACCTGGGCCACAGTGGCCAGTACGGCCAGGGGAGGGACCCATCTGGCTGGAC 413  
Db 241 GCCTTGACCTGGGCCACAGTGGCCAGTACGGCCAGGGGAGGGACCCATCTGGCTGGAC 300  
Qy 414 AATGTGGCTGTGTGGGSCACAGAGAGCTCTTTGGACCAAGTGGGGTCTAATGGCTGGGA 473  
Db 301 AATGTGGCTGTGTGGGSCACAGAGAGCTCTTTGGACCAAGTGGGGTCTAATGGCTGGGA 360  
Qy 474 GTCACTGACTGCACTCAAGAGACGTAGGGGTGATATGACACCCCGGCGGCATCGT 533  
Db 361 GTCACTGACTGCACTCAAGAGACGTAGGGGTGATATGACACCCCGGCGGCATCGT 420  
Qy 534 GGCTACCTTTCTGAACCTGTCTTCCAACTGGCTTGGGCCCCAGGGCTGGAGGAG 593  
Db 421 GGCTACCTTTCTGAACCTGTCTTCCAACTGGCTTGGGCCCCAGGGCTGGAGGAG 480  
Qy 594 GTGGGCTCAAGCCCATCTTGGCCAGTCCAAAGCAGCATAGCCAGTCCAGAGGAGCC 653  
Db 481 GTGGGCTCAAGCCCATCTTGGCCAGTCCAAAGCAGCATAGCCAGTCCAGAGGAGCC 540  
Qy 654 GTGGAGTGAAGTATGAGGGCCACTGGCGCAGGTGTGTGACAGGGCTGGACCATGAAC 713  
Db 541 GTGGAGTGAAGTATGAGGGCCACTGGCGCAGGTGTGTGACAGGGCTGGACCATGAAC 600  
Qy 714 AACAGAGGTGTGTGGGATGTGGGCTTCCAGCAGAGGTGCTGTGACAGCCAC 773  
Db 601 AACAGAGGTGTGTGGGATGTGGGCTTCCAGCAGAGGTGCTGTGACAGCCAC 660  
Qy 774 TACTACAGGAAGTCTGGATCTGAAGTCTGAAGTCTGAAGTCTGAAGTCTGAAGTCTGA 833  
Db 661 TACTACAGGAAGTCTGGATCTGAAGTCTGAAGTCTGAAGTCTGAAGTCTGAAGTCTGA 720  
Qy 834 ACAGATGAAGTCTTCTGATCCACAGGTCACTGCTGGGGACAGAGCCCCACATG 893  
Db 721 ACAGATGAAGTCTTCTGATCCACAGGTCACTGCTGGGGACAGAGCCCCACATG 780  
Qy 894 GCCAATGCCAGGTGCGAGTGGCTCAGCCCGGGGCAAGTGGGCGAGCTGGCCAGGT 953  
Db 781 GCCAATGCCAGGTGCGAGTGGCTCAGCCCGGGGCAAGTGGGCGAGCTGGCCAGGT 840  
Qy 954 GGCATGATGTGTGGTGTGAGTGTGTGGGAGGGCTCACTTCCGCGCCAGCAAGAAAG 1013  
Db 841 GGCATGATGTGTGGTGTGAGTGTGTGGGAGGGCTCACTTCCGCGCCAGCAAGAAAG 900  
Qy 1014 CCACAGCAAGGTCTTGGGACAGAGGCGAGGCTGGGCTGGCTGGGCGGCCAG 1073  
Db 901 CCACAGCAAGGTCTTGGGACAGAGGCGAGGCTGGGCTGGGCTGGGCGGCCAG 960  
Qy 1074 GTGGGCGAGGCGGGTGGAGTGTCTGAACCGCAGTGGGGCAGGGTCTGTGACCAC 1133  
Db 961 GTGGGCGAGGCGGGTGGAGTGTCTGAACCGCAGTGGGGCAGGGTCTGTGACCAC 1020  
Qy 1134 AGTGGAACTCATCTTGGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1193  
Db 1021 AGTGGAACTCATCTTGGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080  
Qy 1194 GAGGCCCTCTTGTGGGCGCGGTGGGCGCAAGGGCTAGGGCCCATCCAGCTGAGTGGGTG 1253

Db 1081 GAGGCCCTCTTGTGGGCGCGGTGGGCGCAAGGGCTAGGGCTATCCACCTGAGTGGGTG 1140  
Qy 1254 CGCTCAGGGGATATGAGCGGACCCCTCAGCACTGCCCTGCTGGGAAGGCTCCAGAAAT 1313  
Db 1141 CGCTCAGGGGATATGAGCGGACCCCTCAGCACTGCCCTGCTGGGAAGGCTCCAGAAAT 1200  
Qy 1314 GGTTCGCAACATGAGAATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1374  
Db 1201 GGTTCGCAACATGAGAATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260  
Qy 1374 AATCAGGTGCGCTTGGCTGGTGGCGGTATCCCTGAGGAGGCGCTATTGGAGGTGCAKGTG 1433  
Db 1261 AATCAGGTGCGCTTGGCTGGTGGCGGTATCCCTGAGGAGGCGCTATTGGAGGTGCAKGTG 1320  
Qy 1434 GAGGTGAACCGGGGTCCACGCTGGGGGAGCGTGTGAGTCAAACTGGGGGCTCACCGAA 1493  
Db 1321 GAGGTGAACCGGGGTCCACGCTGGGGGAGCGTGTGAGTCAAACTGGGGGCTCACCGAA 1380  
Qy 1494 GGCATGTGGCTTGGCGGACAGCTCGGCTGGGTTTTTGGCGATGCCATGCCGTACAGGAAAC 1553  
Db 1381 GGCATGTGGCTTGGCGGACAGCTCGGCTGGGTTTTTGGCGATGCCATGCCGTACAGGAAAC 1440  
Qy 1554 TGGTTCTGTGTGGGACGCCAAGGGGCCAGGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1613  
Db 1441 TGGTTCTGTGTGGGACGCCAAGGGGCCAGGAGTGGTGTGTGTGTGTGTGTGTGTGTGT 1500  
Qy 1614 GGCACAGAGTGGCTGTGAGAGTCCAGAGAGCAGGGCGGTGTGTGTGTGTGTGTGTGTGTGT 1673  
Db 1501 GGCACAGAGTGGCTGTGAGAGTCCAGAGAGCAGGGCGGTGTGTGTGTGTGTGTGTGTGT 1560  
Qy 1674 GCGGGGCTTCTCGCTGGAGTCTCTGATGACAGTCCACAGAGCTGGTGTGTGTGTGTGTGT 1733  
Db 1561 GCGGGGCTTCTCGCTGGAGTCTCTGATGACAGTCCACAGAGCTGGTGTGTGTGTGTGTGT 1620  
Qy 1734 GCGGAGTGTGAGGAGCGGCTACTTGGAGAGCGGCGGTGTGTGTGTGTGTGTGTGTGTGT 1793  
Db 1621 GCGGAGTGTGAGGAGCGGCTACTTGGAGAGCGGCGGTGTGTGTGTGTGTGTGTGTGT 1680  
Qy 1794 GGCACAGGAGAACTGCGCTTCCAGTCTCGGATCAGTGGAGTGGCTGGCTGGCTGGCTGG 1853  
Db 1681 GGCACAGGAGAACTGCGCTTCCAGTCTCGGATCAGTGGAGTGGCTGGCTGGCTGGCTGG 1740  
Qy 1854 GCGGCTTATGCGCTTCTCCACAGAGTCTCAATCTGCGCGAGTGGCTTCTGTGTGTGT 1913  
Db 1741 GCGGCTTATGCGCTTCTCCACAGAGTCTCAATCTGCGCGAGTGGCTTCTGTGTGTGT 1800  
Qy 1914 AAGCTGAGCGGATAGCTGGGTTTGGACCAAGTGGCTGGCTGGCTGGCTGGCTGGCTGG 1973  
Db 1801 AAGCTGAGCGGATAGCTGGGTTTGGACCAAGTGGCTGGCTGGCTGGCTGGCTGGCTGG 1860  
Qy 1974 GTCTTACCCACTACGACCTCTCACTCTCAATGGCTGGCTGGCTGGCTGGCTGGCTGG 2033  
Db 1861 GTCTTACCCACTACGACCTCTCACTCTCAATGGCTGGCTGGCTGGCTGGCTGGCTGG 1920  
Qy 2034 GCGAGCTTCTGTCTGAGGACCAAACTGCGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGG 2093  
Db 1921 GCGAGCTTCTGTCTGAGGACCAAACTGCGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGG 1980  
Qy 2094 GCGAATTTGGAGACAGGGAGTGTGTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2153  
Db 1981 GCGAATTTGGAGACAGGGAGTGTGTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2040  
Qy 2154 GATTGCGAGTGGGTGATATCAGAGATGTGGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGG 2213  
Db 2041 GATTGCGAGTGGGTGATATCAGAGATGTGGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGG 2100  
Qy 2214 GTGAACCCCACTATGAAGTGGAGAGTGTGTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2273  
Db 2101 GTGAACCCCACTATGAAGTGGAGAGTGTGTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2160  
Qy 2274 TGCAAGTATGATGGGCGCGGCTGTGGCTGCACAACTGCTGCACAGGAAATTCATACCCA 2333  
Db 2161 TGCAAGTATGATGGGCGCGGCTGTGGCTGCACAACTGCTGCACAGGAAATTCATACCCA 2220









QY	2052	GACAAAACCTGCCACAGGACTGACAGCGCGGTACCGATGTGCCAACTTTGGAGAACAG	2111	QY	3124	GCAGAAAGACTGCTAGGTAAACACGGACCCAGCCCTGTGAGGGCTCTCTGCCAAGGAAA	3183
Db	2063	GACAAAACCTGCCACAGGACTGACAGCGCGGTACCGATGTGCCAACTTTGGAGAACAG	2122	Db	3126	AAAGAAAGGTGACTGGGGCAGTGTAGACCCAGCTGTGCA-GGTCTCTGCCAGGAAA	3184
QY	2112	GGAGTACTGTAGGTGTGGGACACCTACCGCATGACATTTGATGCGCAGTGGTGGAT	2171	QY	3184	TAATATGGACCACTTTACCTGGCAGGCGAGTCTGCTCTCTCAGGATCACACGCACTCTCA	3243
Db	2123	GGAGTACTGTAGGTGTGGGACACCTACCGCATGACATTTGATGCGCAGTGGTGGAT	2182	Db	3185	AAATCTTGTCTATAATTCAGCAGG-----GAGATCACTGTGCATCTCA	3228
QY	2172	ATCACAGATGTGGGCGCGGAAATATATCTTCCAGGTGATTTGTAACCCCTACTATGAA	2231	QY	3244	GGATTTGCTCTAAACTTCAAGTCTCAACCAAGTGTCTGAAGTGAACCTTTGCATTGAATAA	3303
Db	2183	ATCACAGATGTGGGCGCGGAAATATATCTTCCAGGTGATTTGTAACCCCTACTATGAA	2242	Db	3229	GGATTTGCTCTAAACTTCAAGTCTCAACCAAGTGTCTGAAGTGAACCTTTGCATTGAATAA	3286
QY	2232	CTGCAGAGTCTAGATTTCTTCCAACTATGCTGAGTGGCGCTGCAAGTATGATGGGCAC	2291	QY	3304	TTTTTGGCCATGGGAAAGACATCA----AACAGCCACTCATCTCTACAGAGATAGAAAA	3359
Db	2243	CTGCAGAGTCTAGATTTCTTCCAACTATGCTGAGTGGCGCTGCAAGTATGATGGGCAC	2302	Db	3287	TTTTTGAAGTGGGAAAGACATCAAGTCCACAGGCCACTCATCTCTGCGAGGAAAAAGAGA	3346
QY	2292	GGGTCTGGCTGCACAACTGCCACAGGGAATTCATACCCAGCCCATGCAAGTACTCTCC	2351	QY	3360	CAAGTTTGGCAGAGCAAGACACAGACCCGTGGAGAAATCAGAAAGGGGGAAACAGTCAGT	3419
Db	2303	GGGTCTGGCTGCACAACTGCCACAGGGAATTCATACCCAGCCCATGCAAGTACTCTCC	2362	Db	3347	CAGGTCTAGTAGAGTGA--AGACTAGGCTATGGAGAGTCAAGAGGGGGAGGATTTAAT	3403
QY	2352	CTGCAGAGGAAACAGCGTCTCAGGAAACAACTCTCATCTGAAAGCTGCTCACTGCACTCTA	2411	QY	3420	TTAGTTAAGATGGAACTCTGGGAAAGGCCACCATCTCTGCTTGATGGGCTCTGATTTGC	3479
Db	2363	CTGCAGAGGAAACAGCGTCTCAGGAAACAACTCTCATCTGAAAGCTGCTCACTGCACTCTA	2422	Db	3404	TGAGTTAAGGATTAACCTAGGAAAGGCCACCATCTCTGCTGATGTGGCTGTGATATGC	3463
QY	2412	CTGCTGCGGATACACAGATAGCTCAGCTTATTTGGAGCCATGCTTTCAGAGTCCCA	2471	QY	3480	TCTTGTCTCAAGTGGAAATAAACCCCATGGTCTCTT-GACATGATTTCTTGATCTTTCTC	3538
Db	2423	CTGCTGCGGATACACAGATAGCTCAGCTTATTTGGAGCCATGCTTTCAGAGTCCCA	2482	Db	3464	TCTTGTCTCAAGTGGAAATAAACCCCATGGTCTCTT-GACATGATTTCTTGATCTTTCTC	3523
QY	2472	CTCAGA-GGAAAGGGCCAGTGCCAAAGGGCCACCAAGAACCTGCTCAGGAAGCCTTTTG	2530	QY	3539	CACTGAGACACACTTAAGTATGATCTTTACAGGACTGACACCCCTAATGCCCAATAAAGT	3598
Db	2483	CTCAGA-GGAAAGGGCCAGTGCCAAAGGGCCACCAAGAACCTGCTCAGGAAGCCTTTTG	2542	Db	3524	CACTTGGACACACTGAGAGATGATCTTACAGGCTCTGA(GCTCTCATGCCCAATAAAGT	3583
QY	2531	ATGGCAAGATCACCAATCCAG-ATGGTATTTGCTCCCTCAGGATGCTCTGGGCTCTG-CCC	2588	QY	3599	TGCTCATTTATGGACTGCT	3616
Db	2543	ATAGCAAGATCACCAATCCAG-ATGGTATTTGCTCCCTCAGGATGCTCTGGGCTCTG-CCC	2602	Db	3584	TGCTCATTTATGGACTGCT	3601
QY	2589	CTAAGGCTCTGTGGCTATGGAATATGCTTCCAGGCTTTTGGCTCAGCTGAGTCTCTCTC	2648	RESULT 7			
Db	2603	CTAAGGCTCTGTGGCTATGGAATATGCTTCCAGGCTTTTGGCTCAGCTGAGTCTCTCTC	2662	AF338441			
				LOCUS	AF338441	2271 bp	mrna linear PRI 11-JUL-2001











QY 2220 CCCACTATGAATGGCAGAGTCAGATTTCTCCAAATATGTCGAGTGCGCGTGCAG 2279  
 Db 2101 CCCACTATGAATGGCAGAGTCAGATTTCTCCAAATATGTCGAGTGCGCGTGCAG 2160  
 QY 2280 TATGATGGCAGCGGCTCTGGTGCACACTGCCACACAGGGAATTCATACCCAGCCAAAT 2339  
 Db 2161 TATGATGGCAGCGGCTCTGGTGCACACTGCCACACAGGGAATTCATACCCAGCCAAAT 2220  
 QY 2340 GCAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACCACTCATC 2387  
 Db 2221 GCAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACCACTCATC 2268

RESULT 10  
 AF395336  
 LOCUS AF395336 2271 bp mRNA linear PRI 05-NOV-2001  
 DEFINITION Homo sapiens lysyl oxidase-like 4 (LOXL4) mRNA, complete cds.  
 ACCESSION AF395336  
 VERSION AF395336.1 GI:16660130  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2271)  
 AUTHORS Asuncion,L.P., Fogelgren,B., Fong,K.S.K., Fong,S.F.T., Kim,Y. and  
 Csizsar,K.  
 TITLE A novel human lysyl oxidase-like gene (LOXL4) on chromosome 10q24  
 has an altered SRCR domain and is down-regulated by H-ras  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2271)  
 AUTHORS Asuncion,L.P., Fogelgren,B., Fong,K.S.K., Fong,S.F.T., Kim,Y. and  
 Csizsar,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-JUN-2001) Pacific Biomedical Research Center,  
 University of Hawaii, 1993 East West Road, Honolulu, HI 96822, USA

FEATURES  
 source  
 1..2271  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10q24"  
 /tissue\_type="placenta"  
 gene  
 1..2271  
 /gene="LOXL4"  
 CDS  
 1..2271  
 /gene="LOXL4"  
 /codon\_start=1  
 /product="lysyl oxidase-like 4"  
 /protein\_id="AAL27543.1"  
 /db\_xref="GI:16660131"  
 /translation="MAWSPATLFLFLLLLGQPPRPQSLGTTKRLRVGPSPBEG  
 RLEVHQQGMGTVCDDNFAIQATVACRQIFEAALTWAHSKYGGEGPIWLNVRC  
 VGTSSLQCGNGVSGDSEYGVICHPRHRGYLSETVSNALPGQRLREYVR  
 LKPIILASAKQSPVTEGAVEYKEHWQVCDQGWMTNRSRVCGMLGFFSEVPVDSH  
 YRKVWLKMRDPKRLKSLTKNKSFWIHQVTCLETPHMANQOVAVAPARGKLRPAC  
 PGMHVVSCVAGHPFRPKTKPQKGSNAEBEPRVLRSGAQVGEVRLMNRQGT  
 VCDHWNLSIAVRCQLGFGAREALFGRLCGGLGPIHLSEVRKRGERTLSDCPA  
 LEGSQGCCHENDAAVRCPNPNNGFQNLVLAGGRIPEEGLLEQVVEVNGVPRWSVC  
 SENWGLTEAMQDQRLGLGFAIHAYKETWFSWCTPRAQEVVMGVRCSTGLALQCCQ  
 RHGPVHCSHGGRFLAGVSCMDSAPDLVMNTQLAQETAYLEDRLPLSLYCAHEENCLS  
 KSADHMDPVGYYRLLRFTQIYNLGRTPRPKTGRDSWVHCHQHRHYSLEVETHYD  
 LLTLNGSKVAEGKASFCLEDTNCPGLQRKYACAFGEQVTVGCMQDVRHDDICQW  
 VDITDVGPNGYIFQVIVNPHPEVAESDFSNMLQCRCKYDGRHVRWLHNCHTGNSYPAN  
 AELSLEQRLRNLI"  
 BASE COUNT 454 a 659 c 746 g 412 t  
 ORIGIN

Query Match 62.5%; Score 2258.2; DB 9; Length 2271;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 2263; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 120 ATGGCGTGGTCCCCACAGCCACCTTTCTGTTCTGCTGCTAGGCCAGCCCT 179

1 ATGGCGTGGTCCCCACAGCCACCTTTCTGTTCTGCTGCTAGGCCAGCCCT 60  
 180 CCCAGAGGCCACAGTCACTGGGCACCACTAAGCTCCGCTGCTGGGCCAGAGCAAG 239  
 61 CCCAGAGGCCACAGTCACTGGGCACCACTAAGCTCCGCTGCTGGGCCAGAGCAAG 120  
 240 CCAGAGAGGGCCGCTGGAGGTGCTGCACAGGGCCAGTGGGGCACCTGTGTGATGAC 299  
 121 CCAGAGAGGGCCGCTGGAGGTGCTGCACAGGGCCAGTGGGGCACCTGTGTGATGAC 180  
 300 AACTTTGTATCCAGAGGCCACAGTGGCTTTCGCCAGCTGGGTTCGAAAGTGCCTTG 359  
 181 AACTTTGTATCCAGAGGCCACAGTGGCTTTCGCCAGCTGGGTTCGAAAGTGCCTTG 240  
 360 ACCTGGGCCACAGTGCACAGTGCAGGCCAAGGGAGGAGCCATCTGCTGTCGCAATGTG 419  
 241 ACCTGGGCCACAGTGCACAGTGCAGGCCAAGGGAGGAGCCATCTGCTGTCGCAATGTG 300  
 420 CGCTGTGTGGGCACAGAGAGCTCTTGGACAGTCGCGGTCTAATGGCTGGGGAGTCAGT 479  
 301 CGCTGTGTGGGCACAGAGAGCTCTTGGACAGTCGCGGTCTAATGGCTGGGGAGTCAGT 360  
 480 GACTGCAGTCACTCAGAAAGCTAGGGGTGATATGCCACCCCGCGGCCATCTGCTGGCTAC 539  
 361 GACTGCAGTCACTCAGAAAGCTAGGGGTGATATGCCACCCCGCGGCCATCTGCTGGCTAC 420  
 540 CTTTCTGAAACTGTCTCCAAATGCCCTTGGGCCCGAGGGCCGCTGGAGAGTGCAG 599  
 421 CTTTCTGAAACTGTCTCCAAATGCCCTTGGGCCCGAGGGCCGCTGGAGAGTGCAG 480  
 600 CTCAGGCCATCTTGGCAGTGCAGGCAAGCAGCATACCCAGTGCACGAGGAGCGCTGGAG 659  
 481 CTCAGGCCATCTTGGCAGTGCAGGCAAGCAGCATACCCAGTGCACGAGGAGCGCTGGAG 540  
 660 GTCAAGTATCAGGGCCACTGGCGCAGTGTGTGACAGGGCTGGACCATGAACACAGC 719  
 541 GTCAAGTATCAGGGCCACTGGCGCAGTGTGTGACAGGGCTGGACCATGAACACAGC 600  
 720 AGGCTGTGTGGGATGTCTGGGCTTCCCGACGAGGTGCTGTGCACAGCCACTACTAC 779  
 601 AGGCTGTGTGGGATGTCTGGGCTTCCCGACGAGGTGCTGTGCACAGCCACTACTAC 660  
 780 AGGAACTCTGGGATCTGAAGATGAGGACCCCTAAGTCTAGGCTGAAGAGCTGACGAAT 839  
 661 AGGAACTCTGGGATCTGAAGATGAGGACCCCTAAGTCTAGGCTGAAGAGCTGACGAAT 720  
 840 AAGAACTCTTCTGGATCCACAGTCACTGCTGGGGACAGAGCCCAATGAGTGGCAAC 899  
 721 AAGAACTCTTCTGGATCCACAGTCACTGCTGGGGACAGAGCCCAATGAGTGGCAAC 780  
 900 TGCAGGTGCAGTGGCTCCAGCCCGGGGCAAGTGCAGCCAGCTGCCAGTGGGATG 959  
 781 TGCAGGTGCAGTGGCTCCAGCCCGGGGCAAGTGCAGCCAGCTGCCAGTGGGATG 840  
 960 CATGCTGTGTGCTGCTGTGTGGCAGGGCTCACTTCCGCCCAAGAGCAAGAGCCACAA 1019  
 841 CAGCTGTGTGCTGCTGTGTGGCAGGGCTCACTTCCGCCCAAGAGCAAGAGCCACAA 900  
 1020 CGCAAGAGGTCTCTGGCAGAGAGCCAGGGGTGCGCTCCGCTCCCGGGGCCAGGTGGGC 1079  
 901 CGCAAGAGGTCTCTGGCAGAGAGCCAGGGGTGCGCTCCGCTCCCGGGGCCAGGTGGGC 960  
 1080 GAGGGCCGGGTGGAAGTGTCTCATGAACCCCAAGTGGGGCACCGTCTGTGACCAAGTGG 1139  
 961 GAGGGCCGGGTGGAAGTGTCTCATGAACCCCAAGTGGGGCACCGTCTGTGACCAAGTGG 1020  
 1140 AACTCTATCTCTGCGAGTGTGTGTGCTGCTGAGTGGGCTTTGGCTCTGCTCGGAGGCC 1199  
 1021 AACTCTATCTCTGCGAGTGTGTGTGCTGCTGAGTGGGCTTTGGCTCTGCTCGGAGGCC 1080  
 1200 CTCTTTGGGGCCCGCTGGGGCAAGGGCTAGGGCCATCCACTGACTGAGTGGGCTGC 1259

1081 CTCTTTGGGCGCGCTGGGCAAGGGCTAGGGCCCATCCAGCTGAGTGGGCTGC 1140  
 QY 1260 AGGGATATGAGCGGACCTCAGCACTGCCCTGCCCTGGAAGGGTCCCAAGATGGTTGC 1319  
 Db 1141 AGGGATATGAGCGGACCTCAGCACTGCCCTGCCCTGGAAGGGTCCCAAGATGGTTGC 1200  
 QY 1320 CAACATGAGATGCTGCTGTCTGAGTGCAATGTCCCTAACATGGGCTTTCAGATCAG 1379  
 Db 1201 CAACATGAGATGATGCTGCTGTGAGTGCAATGTCCCTAACATGGGCTTTCAGATCAG 1260  
 QY 1380 GTGGCTTGGCTGGGCGGTATCCCTGAGGAGGGCTATTGGAGGTGCGAGGTGAGGTG 1439  
 Db 1261 GTGGCTTGGCTGGGCGGTATCCCTGAGGAGGGCTATTGGAGGTGCGAGGTGAGGTG 1320  
 QY 1440 AACGGGTGCCACGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1499  
 Db 1321 AACGGGTGCCACGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1380  
 QY 1500 GTGGCTTGGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1559  
 Db 1381 GTGGCTTGGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1440  
 QY 1560 TGGTCCGGGACCCCAAGGGCCAGGAGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1619  
 Db 1441 TGGTCCGGGACCCCAAGGGCCAGGAGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1500  
 QY 1620 GAGTGGGCTTGGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1679  
 Db 1501 GAGTGGGCTTGGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1560  
 QY 1680 CGCTTCCCTGGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1739  
 Db 1561 CGCTTCCCTGGCTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1620  
 QY 1740 CTAGTCCGAGGAGCGGCTTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1799  
 Db 1621 CTAGTCCGAGGAGCGGCTTGGGCGGTGCTGAGTGAAATCGGGGCTCAGCAAGCCATG 1680  
 QY 1800 GAGGAGAACTGCTTCCAAAGTCTGCGGATCAGATGAAATCGGGGCTCAGCAAGCCATG 1859  
 Db 1681 GAGGAGAACTGCTTCCAAAGTCTGCGGATCAGATGAAATCGGGGCTCAGCAAGCCATG 1740  
 QY 1860 CTATTCGCTTCTCCACACAGATCTACAAATCTGGGCGGAGTCTTTCGTCGCAAGACT 1919  
 Db 1741 CTATTCGCTTCTCCACACAGATCTACAAATCTGGGCGGAGTCTTTCGTCGCAAGACT 1800  
 QY 1920 GGACGGATAGCTGGGTTTGGGCGGATGCGGCGGATGCGGCGGATGCGGCGGATGCGGCGG 1979  
 Db 1801 GGACGGATAGCTGGGTTTGGGCGGATGCGGCGGATGCGGCGGATGCGGCGGATGCGGCGG 1860  
 QY 1980 ACCACTACGACCTCTCTCAATGGCTCCAAAGTGGCTGAGGGGCGGCGGCGGCGGCGGCGG 2039  
 Db 1861 ACCACTACGACCTCTCTCAATGGCTCCAAAGTGGCTGAGGGGCGGCGGCGGCGGCGGCGG 1920  
 QY 2040 TTCTGTCTGAGGACACAAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2099  
 Db 1921 TTCTGTCTGAGGACACAAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1980  
 QY 2100 TTGGAGAACAGGAGTGAATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2159  
 Db 1981 TTGGAGAACAGGAGTGAATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2040  
 QY 2160 CAGTGGGTGATATCACAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2219  
 Db 2041 CAGTGGGTGATATCACAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2100  
 QY 2220 CCCACTATGAGTGGGAGTGAATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2279  
 Db 2101 CCCACTATGAGTGGGAGTGAATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2160  
 QY 2280 TATGATGGGACCGGGTCTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2339  
 Db 2161 TATGATGGGACCGGGTCTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2220

QY 2340 GCAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACAACCTCATCTGA 2390  
 Db 2221 GCAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACAACCTCATCTGA 2271

RESULT 11  
 AK025542 Homo sapiens cDNA: FLJ21889 fis, clone HEP03178. linear  
 LOCUS AK025542.1 GI:10438090  
 DEFINITION oligo capping; fis (full insert sequence).  
 ACCESSION Homo sapiens hepatoma cell\_line HepG2 cDNA to mRNA, clone lib:HEP  
 KEYWORDS clone:HEP03178.  
 SOURCE

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS  
 Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,  
 Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,  
 Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2059)  
 Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,  
 Shibahara, T., Tanaka, T. and Nakamura, Y.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,  
 University of Tokyo, Laboratory of Genome Structure Analysis, Human  
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
 Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
 Fax:81-3-5449-5416)  
 NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan: cDNA full insert  
 sequencing: Research Association for Biotechnology: cDNA library  
 construction, 5'- and 3'-end one pass sequencing: Department of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).

COMMENT  
 Location/Qualifiers  
 1..2059  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HEP03178"  
 /cell\_line="HepG2"  
 /clone\_lib="HEP"  
 /note="cloning vector pME18-IL3"  
 3..800  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB15167.1"  
 /db\_xref="GI:10438091"  
 /translation="MSGVRCSTELALQCORHGPVHCSSKGGKFLAGVSCMSAPDL  
 VMNQLQETAYLEDRPLSOLYCAHEENFLSKSADHMDWPGYVRLRLRFTQIYNLAR  
 TDFPKTGRDSWVWHQCHRHVHSIEVETFDLLTLNGSKVAGHKAFCELEDNCPPTG  
 LORRYACANFGQVTCGMDTYRHDIIDTOMVDITDVGNGYIPQIVNPHVEAESD  
 FSNMLQCRQYDGHVRLHNCHTGNSYPANAEISLEOEORLNNLI"

BASE COUNT 544 a 547 c 511 g 457 t  
 ORIGIN  
 Query Match 55.9%; Score 2019.6; DB 9; Length 2059;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2022; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CDS  
 1591 TGATGAGTGGGTGGCTGCTCAGGCACAGAGCTGGCCCTGCAGCAGTGGCAGGCGACG 1650  
 1 TGATGAGTGGGTGGCTGCTCAGGCACAGAGCTGGCCCTGCAGCAGTGGCAGGCGACG 60  
 1651 GGCCGCTGCACTGCTCCCAAGCGTGGCGGGCGCTTCCTGCTGGAGTCTCTTCATGGACA 1710  
 61 GGCCGCTGCACTGCTCCCAAGCGTGGCGGGCGCTTCCTGCTGGAGTCTCTTCATGGACA 120

QY 1711 GTGCAACGAGACCTGGTGAAGAACGCCAGCTAGTGCAGGAGACGCCCTACTTTGGAGGACC 1770  
Db 121 GTGCAACGAGACCTGGTGAAGAACGCCAGCTAGTGCAGGAGACGCCCTACTTTGGAGGACC 180  
QY 1771 GCCCGCTCAGCAGCTGTATTGTGCCACGAGAGAACTGCTCTCCAAGTCTGGCGATC 1830  
Db 181 GCCCGCTCAGCAGCTGTATTGTGCCACGAGAGAACTGCTCTCCAAGTCTGGCAGATC 240  
QY 1831 ACATGAGCTGGCCCTCAGGATACCGCCCTATTGCGCTTCTCCACACAGATCTCAATC 1890  
Db 241 ACATGAGCTGGCCCTCAGGATACCGCCCTATTGCGCTTCTCCACACAGATCTCAATC 300  
QY 1891 TGGGCGGAGCTGACTTTCGTCCAAAGACTGGACCGGATAGTGGGTTTGGCACCAAGTGCC 1950  
Db 301 TGGGCGGAGCTGACTTTCGTCCAAAGACTGGACCGGATAGTGGGTTTGGCACCAAGTGCC 360  
QY 1951 ACAGGCAATPACACAGCAATGAGGCTTTCACCCACTACGACTCTCTCAATGGCT 2010  
Db 361 ACAGGCAATPACACAGCAATGAGGCTTTCACCCACTACGACTCTCTCAATGGCT 420  
QY 2011 CCAAGGTGCTGAGGGGCAACAGGCGAGCTTCTGTCTGGAGGACACAACTGCCCCCAG 2070  
Db 421 CCAAGGTGCTGAGGGGCAACAGGCGAGCTTCTGTCTGGAGGACACAACTGCCCCCAG 480  
QY 2071 GACTGACGGCGCTACGATGTGCCAACTTTGGAGAACAGGAGTGAAGTGTGCT 2130  
Db 481 GACTGACGGCGCTACGATGTGCCAACTTTGGAGAACAGGAGTGAAGTGTGCT 540  
QY 2131 GGGACACCTACCGGATGACATTGATTCGCAAGTGGGATGATCACAGATGCGGCCCG 2190  
Db 541 GGGACACCTACCGGATGACATTGATTCGCAAGTGGGATGATCACAGATGCGGCCCG 600  
QY 2191 GGAATTATATCTTCCAGGTGATTTGAACCCCACTATGAAGTGGCAGAGTCAGATTTCT 2250  
Db 601 GGAATTATATCTTCCAGGTGATTTGAACCCCACTATGAAGTGGCAGAGTCAGATTTCT 560  
QY 2251 CCAACAATATGTGAGTGCCTGCAAGTATGATGGCACCGGCTTGGCTGCACACT 2310  
Db 661 CCAACAATATGTGAGTGCCTGCAAGTATGATGGCACCGGCTTGGCTGCACACT 720  
QY 2311 GCCACACAGGAAATTCATACCCAGCCAAATGGAGAACTCTCCCTGGAGCAGGAAAGCGTC 2370  
Db 721 GCCACACAGGAAATTCATACCCAGCCAAATGGAGAACTCTCCCTGGAGCAGGAAAGCGTC 780  
QY 2371 TCAGGAACAACCTCATCTGAAGTGTCACTGCACACTCTCTGCTCCCGATACACAG 2430  
Db 781 TCAGGAACAACCTCATCTGAAGTGTCACTGCACACTCTCTGCTCCCGATACACAG 840  
QY 2431 ATACCTCAGCTTATTGGAGCATGCCCCTTCACAGAGTCCCAACTCAGAGGAAAGGGCCA 2490  
Db 841 ATACCTCAGCTTATTGGAGCATGCCCCTTCACAGAGTCCCAACTCAGAGGAAAGGGCCA 900  
QY 2491 GTSCCAAGGGGACCAAGAACCTGCTCAGGAAGCCTTTTGTGGCAAGATCACCAATCCA 2550  
Db 901 GTSCCAAGGGGACCAAGAACCTGCTCAGGAAGCCTTTTGTGGCAAGATCACCAATCCG 960  
QY 2551 GATGTATTGCTCCCTCAGGATGGCTCTGGCCCTGCGCCCTAAGGCGCTGTGGCCTATGGA 2610  
Db 961 GATGTATTGCTCCCTCAGGATGGCTCTGGCCCTGCGCCCTAAGGCGCTGTGGCCTATGGA 1020  
QY 2611 ATATGCTCCAGGCTTTGCTCAGCTGAGTCTCTTCTGTGAAGAAACCCAGTATCCC 2670  
Db 1021 ATATGCTCCAGGCTTTGCTCAGCTGAGTCTCTTCTGTGAAGAAACCCAGTATCCC 1080  
QY 2671 TGAATCTTCCACAGAGATCCGGGATTCAGAGCTCTCAGTTCTTAGGGATGGACTATG 2730  
Db 1081 TGAATCTTCCACAGAGATCCGGGATTCAGAGCTCTCAGTTCTTAGGGATGGACTATG 1140  
QY 2731 GCCCAGTCCCCCACTCAAGTGGTGTCTTGGAAATGTCTTGGAGGAGTATAGACAGAGGA 2790  
Db 1141 GCCCAGTCCCCCACTCAAGTGGTGTCTTGGAAATGTCTTGGAGGAGTATAGACAGAGGA 1200

QY 2791 CCAAAATACACAGAGGTAGTGTAGTCTCTGTAGGAGCTCAAAAGCAACAACTTGT 2850  
Db 1201 CCAAAATACACAGAGGTAGTGTAGTCTCTGTAGGAGCTCAAAAGCAACAACTTGT 1260  
QY 2851 ATCAAAATCAAACTGGCAGAGAGCTGGTGGATCCAATCCTTCTCATCTGTGTAT 2910  
Db 1261 ATCAAAATCAAACTGGCAGAGAGCTGGTGGATCCAATCCTTCTCATCTGTGTAT 1320  
QY 2911 TTGAATCTCAGCTCTCACACTCTGTCTTTAGTCTCTTACCTTTATCTTACACACACA 2970  
Db 1321 TTGAATCTCAGCTCTCACACTCTGTCTTTAGTCTCTTACCTTTATCTTACACACACA 1380  
QY 2971 TGGGTGTTTCTTATTTATCTTGAAGCAGACAGCTCGGGCATCCCTTTATTCCTGTATGG 3030  
Db 1381 TGGGTGTTTCTTATTTATCTTGAAGCAGACAGCTCGGGCATCCCTTTATTCCTGTATGG 1440  
QY 3031 CCAACACCAACAGTTACGAGTGTGTGAGAGGGGCAAGTTTTCACAGAAATGGCCAGATA 3090  
Db 1441 CCAACACCAACAGTTACGAGTGTGTGAGAGGGGCAAGTTTTCACAGAAATGGCCAGATA 1500  
QY 3091 GGGCCTTCTTACAGAGCAGCAAGAGTAGGCCAAGCAGAAAGTCTGCTGAGGTAAACGGA 1560  
Db 1501 GGGCCTTCTTACAGAGCAGCAAGAGTAGGCCAAGCAGAAAGTCTGCTGAGGTAAACGGA 1620  
QY 3151 CCCAGCCCTGTCTCAGGCTCTGCCAAGGAAATATATGGACCAATTTACCTGGCAGGCA 3210  
Db 1561 CCCAGCCCTGTCTCAGGCTCTGCCAAGGAAATATATGGACCAATTTACCTGGCAGGCA 1620  
QY 3211 GTCTGCTCTCTCAGGATCACCGCATCTCAGGATTCGTCTAAACTTCAAGTCTCAAC 3270  
Db 1621 GTCTGCTCTCTCAGGATCACCGCATCTCAGGATTCGTCTAAACTTCAAGTCTCAAC 1680  
QY 3271 CAAGTGTCTGAAGTGAACTTTGCAATTTGATTAATTTTCCCATGGAAAGCAATCAACA 3330  
Db 1681 CAAGTGTCTGAAGTGAACTTTGCAATTTGATTAATTTTCCCATGGAAAGCAATCAACA 1740  
QY 3331 AGGCATCTATCTTACAGAGATAAGAAACAAAGTTTGGCAGAGCAAGAGACAGAACCG 3390  
Db 1741 AGGCATCTATCTTACAGAGATAAGAAACAAAGTTTGGCAGAGCAAGAGACAGAACCG 1800  
QY 3391 TGGAGAAATCAGAGAGGGGAAACAGTCAGTTTGTAGTAAAGATGSAACCTGGAAAGGCCAC 3450  
Db 1801 TGGAGAAATCAGAGAGGGGAAACAGTCAGTTTGTAGTAAAGATGSAACCTGGAAAGGCCAC 1860  
QY 3451 CATTCCTGTTGATGGGCTCTGATTTGCTCTTGTCTCAAGTGGAAATAAAACCCCATGGTC 3510  
Db 1861 CATTCCTGTTGATGGGCTCTGATTTGCTCTTGTCTCAAGTGGAAATAAAACCCCATGGTC 1920  
QY 3511 TTCTTGACATGATTTCTGATCTTTTCTCCACTGAGACACACTTAAAGTATGATCCCTTACA 3570  
Db 1921 TTCTTGACATGATTTCTGATCTTTTCTCCACTGAGACACACTTAAAGTATGATCCCTTACA 1980  
QY 3571 GGACTGACCCCTAATGCCAATAAAAGTTGCTCATTTATGGACTGCT 3616  
Db 1981 GGACTGACCCCTAATGCCAATAAAAGTTGCTCATTTATGGACTGCT 2026

RESULT 12  
AF338440

LOCUS

DEFINITION

Mus musculus

cds.

ACCESSION

AF338440

VERSION

AF338440.1

GI:14669468

KEYWORDS

SOURCE

Mus musculus.

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 2274)

AUTHORS

Ito, H., Akiyama, H., Iguchi, H., Iyama, K., Miyamoto, M., Ohsawa, K. and Nakamura, I.

Molecular cloning and biological activity of a novel lysyl

AF338440 2274 bp mRNA linear ROD 11-JUL-2001  
Mus musculus lysyl oxidase-related protein C (LoxC) mRNA, complete

cds.

ACCESSION

AF338440

VERSION

AF338440.1

GI:14669468

KEYWORDS

SOURCE

Mus musculus.

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 2274)

AUTHORS

Ito, H., Akiyama, H., Iguchi, H., Iyama, K., Miyamoto, M., Ohsawa, K. and Nakamura, I.

Molecular cloning and biological activity of a novel lysyl

oxidase-related gene expressed in cartilage  
 J. Biol. Chem. 276 (26), 24023-24029 (2001)  
 21316447  
 PUBMED  
 REFERENCE 2 (bases 1 to 2274)  
 Akiyama, H., Ito, H. and Nakamura, T.  
 Direct Submission  
 Submitted (19 JAN 2001) Department of Orthopaedics, Kyoto  
 University, 54 Shogoin-Kawahara-cho, Sakyo, Kyoto 606-8507, Japan  
 JOURNAL  
 TITLE Location/Qualifiers  
 FEATURES  
 source  
 1..2274  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_line="ATDC5"  
 /cell\_type="chondrogenic EC"  
 1..2274  
 /gene="LoxC"  
 1..2274  
 /gene="LoxC"  
 1..2274  
 /note="collagen cross-link"  
 /codon\_start=1  
 /product="lysyl oxidase-related protein C"  
 /protein\_id="AAK71933.1"  
 /db\_xref="GI:14669469"  
 /translation="MMWPPPTFLFLLLLSQAPSRPOSSGKTLRLVGPADPPEE  
 GRLEVLHOGMTGCDDEDFALQPAETVACRLQGFESALTWAHSKYGOGEGPIWLDNVR  
 CLGTEKTLDOCSNGMIGSDCHSEDDVGVCHPRQGHVHSEKVSNALGPQGRLEEV  
 RLKPLASAKRHSPTGAEVRYDGHVQVCDQMTMNSRVVUCMLGPFSQTSVNS  
 HYKRVNLMKDKPRLSLTKNSFWIHRVDCFGTEPHLAKRCQVAVPGRKLKRP  
 CPGMHAVSVVAGPHFRKPKRKEKSHAEELKRLRSGAOGVEGRVFLMNRQWG  
 TVDHRWNLISASVVCRLQFGSAREALFQAQLGQLGPIHLSEVRCRGVETLGDCL  
 ALESGONGCHANDAAVRNIIDMGFQNKVRLAGRNSEBVEVVEVANGGPRWGT  
 CSIDHMLTEAMVTCRLQGLGFANFALKTWYWGQTEPEAKVNVMSGVRCSGTGEMALQCC  
 QHGPVTHCSHGPRFSAGVACMNSAPLVNNAOLVOETAYLEDRPLSLYCAHEENCL  
 SKSDHMDPKYGRRLRRESSOYLNLGRADFRPKGRHSWLMHCHRNHSLIEVFTHY  
 DULLTNGSKVAGHAKSFCLEBETNCSGVQRRYACANFQGVAVGWMDFYVHDIQO  
 WADITDVGPDYTFQVYVNPNTINDVAESDFSNMIRCKRYDQGVWLNCHTGDYSYRA  
 NAEISLEQEQRLRNLI"  
 BASE COUNT 495 a 620 c 715 g 444 t  
 ORIGIN  
 Query Match 47.8%; Score 1730; DB 10; Length 2274;  
 Best Local Similarity 85.4%; Pred. No. 0;  
 Matches 194; Conservative 0; Mismatches 330; Indels 3; Gaps 1;  
 QY 120 ATGGCGTGGTCCCGACAGCCACCTCTTTCTGTTCTG---CTGCTGCTAGGCCAGCC 176  
 DB 1 ATGATGTGGCCCAACCCACCTCTTCCCTGTTCTGCTACTGCTGCTAAAGCCAGCC 60  
 QY 177 CTTCCAGCAGCCACAGTCACCTGGGACCACTAAGCTCCGGCTGGTGGGCCAGAGAGC 236  
 DB 61 CCTTCCAGTAGCCACAGTCATCAGCACCAAGAGCTCAGGCTTGTGGGCCAGCGGAC 120  
 QY 237 AAGCCAGAGAGGCGCGCTGGAGGTGCTGCACAGGGCCAGTGGGGACCGTGTGTGAT 296  
 DB 121 AGACCAGAGAGGCGCGCTTGGAGGTGCTGCACAGGGCCAGTGGGGACCGTGTGTGAT 180  
 QY 297 GACAACTTGTGATCCAGGAGGCGACAGTGGCTTGGCCGACAGCTGGGCTTGAAGCTGCC 356  
 DB 181 GATGATTTCCGCTCTCCAGGAGGCTACTGTGGCTTGGCCGACAGCTGGGCTTGAAGTCAGCC 240  
 QY 357 TTGACTGGGCGCACAGTCCAGTAGTCCGACAGGAGGAGGACCATCTGGCTGGACAAT 416  
 DB 241 TTGACTGGGCGCACAGTCCAGTAGTGGTCAAGAGAGGAGTCCCATCTGGCTGGACAAT 300  
 QY 417 GTGCGCTGTGTGGGCGACAGAGCTCTTGGACAGTCCGGGTCTTAATGGCTGGCGAGTC 476  
 DB 301 GTTCGTTGTTGGGCGACAGAGAGCCCTAGATCAGTGTGGCTCTTAACGGCTGGGGTATC 360  
 QY 477 AGTGACTGCAGTCACTCAGAAGACGTAGGGGTGATATGCCACCCCGCGGCATCGTGGC 536  
 DB 361 AGTGACTGCAGTCACTCAGAGATGTTGGGGTGTGTATGTACCCACCGCGCCAGCAGGA 420

537 TACCTTTCTGAACCTGTCTCAATGCCCTTGGGCCCCAGGCCGCGCTGGAGGAGGTG 596  
 DB 421 TATCACTCTGAGAAGGTCTCAATGCCCTCGGGCTCAGGCCGCGCTAGAGAGGTA 480  
 QY 597 CGGCTCAAGCCCATCTTGGCAGTGGCAAGCAGCATAGCTCAGTCAACCGAGGAGCGGTG 656  
 DB 481 CGGCTCAAGCCCATCTTGGCAGTGGCAAGCAGCATAGCTCAGTCAACCGAGGAGGTG 540  
 QY 657 GAGTGAAGTATGAGGGCCCACTGGCGCAGGTGTGTGACAGGGCTGGACCATGAACAAC 716  
 DB 541 GAAGTACGTTACAGCGCCACTGAGGAGCAGGTGTGTGACAGGGCTGGACCATGAACAAC 600  
 QY 717 AGCAGGTGTGTGGGATGTGGGCTTCCCGAGGAGGTGCTGTGTGACAGGCTGACATAC 776  
 DB 601 AGCAGGTGTGTATCGGGGATGTGGGCTTCCCGAGTCAATCTGTGTGACAGGCTGACATAC 660  
 QY 777 TACAGGAAGTCTGGGATCTCAACATGAGGACCTTAAG1CTAGGCTGAAGAGCTTGACG 836  
 DB 661 TACAGGAAGTCTGGGATCTCAACATGAGGATCCCAAG1CTAGGCTCAACAGCTGACA 720  
 QY 837 AATAAGAACTCTTCTGGATCCCAAGGTCACTGCTGCTGGGACAGAGCTTACATGGCC 896  
 DB 721 AAAAAGAACTCTTCTGGATTCACCGGTTGACTGTTTCCGGACAGAGCTTACATGGCC 780  
 QY 897 AACTGCCAGGTGAGGTGGCTTCCAGCCCGGGGCAAGCTGCGGCCAGCTTCCCGAGCTGGC 956  
 DB 781 AAGTGCCAGGTACAGGTGGCTTCCAGGAGGAGGCAAGCTTCCCGAGCTTCCAGCGCCG 840  
 QY 957 ATSCATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1016  
 DB 841 ATGACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 1017 CAACGCAAGAGTCTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1076  
 DB 901 ACAGCGAAGAGTCTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 QY 1077 GCGAGGCGCGGTGGAGTGTCTATGAACCGCAGTGGCGACCGTCTGTGACACAGG 1136  
 DB 961 GGTGAGGCGCGGTGGAGTGTCTATGAACCGCAGTGGCGACCGTCTGTGACACAGG 1020  
 QY 1137 TGAACCTCATCTTCCAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1196  
 DB 1021 TGAACCTCATCTTCCAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080  
 QY 1197 GCCTCTTTGGGCGCGGTGGGCGCAGGAGTGGGCGCCTCCACCTGAGTGGAGTGGCG 1256  
 DB 1081 GCCCTTTTGGGCGCGGTGGGCGCAGGAGTGGGCGCCTCCACCTGAGTGGAGTGGCG 1140  
 QY 1257 TGCAGGCGATATGAGCGGACCTCAGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1316  
 DB 1141 TGCAGGCGATATGAGCGGACCTCAGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 QY 1317 TGCACATCAGATGT 1376  
 DB 1201 TGTCAACATCAACAGCGT 1260  
 QY 1377 CAGGTGCGT 1436  
 DB 1261 AAGTGTGCGT 1320  
 QY 1437 GTCAACGGGTCCACAGCTGGGAGCGT 1496  
 DB 1321 GTGATGTGGGTCCACGATGGGAGT 1380  
 QY 1497 ATGTGCGCTGCGCAGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1556  
 DB 1381 ATGTGCGCTGCGCAGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1440  
 QY 1557 TTCTGTGCGGAGCGCCAAAGGCGCCAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1616  
 DB 1441 TACTGGCAGGGGACACAGAGGCGCCAAAGAGT 1500  
 QY 1617 ACAGAGTGTGCGCTGCGCAGAGTGTGCGCAGAGGCGCCAGGCGCGTGTGCTGCTGCTGCTGCTGCTG 1676



Db 1501 ACAGAAATGGCCCTGACAGAGTGTACAGAGCATGGGCGGTGCACTGTCCACGGGCCA 1560  
 QY 1677 GGGCGCTTCTCGCTGGAGTCTCTCATGACAGTGCACAGACCTGGTGTATGAACGCC 1736  
 Db 1561 GGGCGCTTCTCGCTGGAGTCTCTCATGACAGTGCACAGACCTGGTGTATGAACGCC 1620  
 QY 1737 CAGCTAGTGCAGAGACGGCCTACTTGGAGGACCGCCGCTCAGCCAGCTGTATTGTGCC 1796  
 Db 1621 CAGCTGTACAGAGACGGCCTACTTGGAGGATCGTCCACTCAGCATGCTGTACTGTCT 1680  
 QY 1797 CACGAGGAACTGCTCTTCCAGTGTGGGATCACATGACATGGCCCTACGGATACCGC 1856  
 Db 1681 CACGAGGAAACTGCTCTTCCAGTGTGGGATCACATGACATGGCCCTACGGATACCGC 1740  
 QY 1857 GCGCTATTGCGCTTCTCCACAGATCTACAATCTGGGCGGAGTACTTTCGTCACAG 1916  
 Db 1741 GCGTTGCTGCGCTTCTCTCCACAGATCTACAATCTGGGCGGAGTACTTTCGTCACAG 1800  
 QY 1917 ACTGACGCGATAGTGGGTTTGGCCAGTGCACAGGATTAACACAGCATTTGAGGTC 1976  
 Db 1801 GCTGACGCGATAGTGGGTTTGGCCAGTGCACAGGATTAACACAGCATTTGAGGTC 1860  
 QY 1977 TTCACCCACTAGACCTCTCTCACTCTCAATGGGTCTCCAGGTGGTGGAGGCAACAAGGCC 2036  
 Db 1861 TTCACCTATTAGCTCTCTCACTCTCAATGGGTCTCCAGGTGGTGGAGGCAACAAGGCC 1920  
 QY 2037 AGCTTCTGTCTGAGGACACAACTGCCCCACAGCATGTCAGCGCGCTAGCATGTGCC 2096  
 Db 1921 AGCTTCTGTCTGAGGACACAACTGCCCCACAGCATGTCAGCGCGCTAGCATGTGCC 1980  
 QY 2097 AACTTTGGAGAACAGGAGTGTAGTGGTGTGGGACACCTACCGGCATGACATTCAT 2156  
 Db 1981 AACTTTGGAGAACAGGAGTGTAGTGGTGTGGGACACCTACCGGCATGACATTCAT 2040  
 QY 2157 TGCCAGTGGTGGATATCAGATGTGGGCGCGGGAATATATCTTCCAGGTGATTGTG 2216  
 Db 2041 TGCCAGTGGTGGATATCAGATGTGGGCGCGGGAATATATCTTCCAGGTGATTGTG 2100  
 QY 2217 AACCCCCACTATGAGTGGCAGATGATTTCTCCAAATATGCTGCAGTCCGCTGC 2276  
 Db 2101 AACCCCAACAGATTTGAGAGTTCGATTTCTCCAAATATGATACGATGCGCTGC 2160  
 QY 2277 AAGTATGATGGGACCGGCTCTGGCTGCACAACTGCCACACAGGGAATTCATCCAGCC 2336  
 Db 2161 AAGTATGATGGGACCGGCTCTGGCTGCACAACTGCCACACAGGGAATTCATCCAGCC 2220  
 QY 2337 AATGAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACAACCTCATCTGA 2390  
 Db 2221 AATGAGAACTCTCCCTGGAGCAGGAACAGCGTCTCAGGAACAACCTCATCTGA 2274

RESULT 13  
 AL139241/c 18912 bp DNA linear PRI 26-OCT-2001  
 LOCUS Human DNA sequence from clone RP11-34A14 on chromosome 10, complete  
 DEFINITION sequence.

ACCESSION AL139241  
 VERSION AL139241.11 GI:16508186  
 KEYWORDS HTG

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 18912)  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE Direct Submission  
 JOURNAL Submitted (26-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk

COMMENT On Oct 29, 2001 this sequence version replaced gi:14575095.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., paired quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr10  
 RP11-34A14 is from the library RP11-11.1 constructed by the group of Peter de Jong. For further details see

http://www.chori.org/bacpac/home.htm  
 VECTOR: pBACe3.6  
 This sequence is the entire insert of clone RP11-34A14. The true left end of clone RP11-439D8 is at 146755 in this sequence. The true right end of clone RP11-594J24 is at 52695 in this sequence.

## FEATURES

## source

1. 188192  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

/chromosome="10"  
 /clone="RP11-34A14"  
 /clone\_lib="RP11-11.1"

## misc\_feature

20413..20507

/note="Sequence from overlapping clone  
 RP11-594J24 (AL355301). Assembly confirmed by restriction digest."

## misc\_feature

101262..101305

/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

## BASE COUNT

49922 a 40190 c 42007 g 56073 t

## ORIGIN

Query Match 36.1%; Score 1303.8; DB 9; Length 188192;  
 Best Local Similarity 99.8%; Pred. No. 8.3e-296;  
 Matches 1305; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2310 TGGCACACAGGGAATTCATACCCAGCCCAATGCAGAACTCTCCCTGGAGCAGGAACAGCGT 2369  
 Db 81880 TGGCCACAGGGAATTCATACCCAGCCCAATGCAGAACTCTCCCTGGAGCAGGAACAGCGT 81821  
 QY 2370 CTCAGGAACAACCTCATCTGAAGCTGTCACTGCACACTCTTAGCTGCTGCCGATACACCA 2429  
 Db 81820 CTCAGGAACAACCTCATCTGAAGCTGTCACTGCACACTCTTAGCTGCTGCCGATACACCA 81761  
 QY 2430 GATACCTCAGCTTATTGAGCCATGCCCTTTCACAGAGTCCCAACTCAGAGGAAAGGGCC 2489  
 Db 81760 GATACCTCAGCTTATTGAGCCATGCCCTTTCACAGAGTCCCAACTCAGAGGAAAGGGCC 81701  
 QY 2490 AGTGCCAAAGGGGACCAAGAACCTGCTCAGGAAGCCCTTTTGATGGCAAGATCACCATCC 2549  
 Db 81700 AGTGCCAAAGGGGACCAAGAACCTGCTCAGGAAGCCCTTTTGATGGCAAGATCACCATCC 81641  
 QY 2550 AGATGGTATTGCTCCCTCAGGATGGCTTGGGCTGCCCTTAGGCGCTGTGGCTATGG 2609  
 Db 81640 AGATGGTATTGCTCCCTCAGGATGGCTTGGGCTGCCCTTAGGCGCTGTGGCTATGG 81581  
 QY 2610 AATATGCTCCAGGCTTTGGCTCAGCTGAGCTCCTCTTCTGTAAGGAAACCCAGTCAATCC 2669  
 Db 81580 AATATGCTCCAGGCTTTGGCTCAGCTGAGCTCCTCTTCTGTAAGGAAACCCAGTCAATCC 81521

QY 2670 CTGAATCTTGGCACAGAGATCCGGATTCCAGGAGCTCTCAGTTTCTTAGGATGGACTAT 2729  
 DB 81520 CTGAATCTTGGCACAGAGATCCGGATTCCAGGAGCTCTCAGTTTCTTAGGATGGACTAT 81461  
 QY 2730 GGGCCAGTCCGCCATCTAAGTGGTGGTTCGAAATCTCTGGAGGAGTATAGGACAGG 2789  
 DB 81460 GGGCCAGTCCGCCATCTAAGTGGTGGTTCGAAATCTCTGGAGGAGTATAGGACAGG 81401  
 QY 2790 ACCAAATACACAGCAGGAGTGTAGTCTCTGTAGGAGCTCAAGACACAACTTG 2849  
 DB 81400 ACCAAATACACAGCAGGAGTGTAGTCTCTGTAGGAGCTCAAGACACAACTTG 81341  
 QY 2850 TATCAAAATCAAACTGGCAGAGAAGCTGGTGGATCAATCTCTTCATCTGCTTGA 2909  
 DB 81340 TATCAAAATCAAACTGGCAGAGAAGCTGGTGGATCAATCTCTTCATCTGCTTGA 81281  
 QY 2910 TTTAGAACTCACTCTCACACTCTGTCTTTAGTGTCTTACCTTTATCTTACACACAC 2969  
 DB 81280 TTTAGAACTCACTCTCACACTCTGTCTTTAGTGTCTTACCTTTATCTTACACACAC 81221  
 QY 2970 ATGGTGTCTTATCTTATCTTGAAGCAGACAGCTCGGATCCCTTATGCTGATGG 3029  
 DB 81220 ATGGTGTCTTATCTTATCTTGAAGCAGACAGCTCGGATCCCTTATGCTGATGG 81161  
 QY 3030 GCCAACACCAAGTTACCGAGTGTCTTGAAGGGGCAAGTTTACAGAAATGGCCAGAT 3089  
 DB 81160 GCCAACACCAAGTTACCGAGTGTCTTGAAGGGGCAAGTTTACAGAAATGGCCAGAT 81101  
 QY 3090 AGGCGCTTCTACAGAGCAGAGTGTAGGCAAGCAGAACTGTCTGAGTTAAGCAGG 3149  
 DB 81100 AGGCGCTTCTACAGAGCAGAGTGTAGGCAAGCAGAACTGTCTGAGTTAAGCAGG 81041  
 QY 3150 ACCCAGCCCTCTCAGGCGCTCTGCAAGGAATATATGGACCATTTACCTGGCAGGC 3209  
 DB 81040 ACCCAGCCCTCTCAGGCGCTCTGCAAGGAATATATGGACCATTTACCTGGCAGGC 80981  
 QY 3210 AGTGTGTCTCTCAGAGTATCCAGCATCTCAGATTTGGTCTAAATCTCAAGTCTCAA 3269  
 DB 80980 AGTGTGTCTCTCAGAGTATCCAGCATCTCAGATTTGGTCTAAATCTCAAGTCTCAA 80921  
 QY 3270 CCAAGTGTCTGAAGTGAATCTTGATTTGAATTAATTTTCCATGGAAGAACATCAAC 3329  
 DB 80920 CCAAGTGTCTGAAGTGAATCTTGATTTGAATTAATTTTCCATGGAAGAACATCAAC 80861  
 QY 3330 AAGCCACTCATCTACAGAGATAAGAAACAAGTTTGGCAGAGCAGACAGACAGACC 3389  
 DB 80860 AAGCCACTCATCTACAGAGATAAGAAACAAGTTTGGCAGAGCAGACAGACAGACC 80801  
 QY 3390 GTGGAGAAATCAGAGGGGCAAGTCAAGTTTAAAGTGAACCTGGCAAGGCCA 3449  
 DB 80800 GTGGAGAAATCAGAGGGGCAAGTCAAGTTTAAAGTGAACCTGGCAAGGCCA 80741  
 QY 3450 CCATTCCTGCTGTAGTGGGCTCTGATTTGCTCTGCTCAAGTGAATAAACCCTATGGT 3509  
 DB 80740 CCATTCCTGCTGTAGTGGGCTCTGATTTGCTCTGCTCAAGTGAATAAACCCTATGGT 80681  
 QY 3510 CTTCTTGACATCTTCTGATCTTTCTCCACAGACACACTTAAAGTGAATGATCTTAC 3569  
 DB 80680 CTTCTTGACATCTTCTGATCTTTCTCCACAGACACACTTAAAGTGAATGATCTTAC 80621  
 QY 3570 AGGACTGACACCTTAATGCCAATAAAGTTGCTCATTTATGGACTGCT 3616  
 DB 80620 AGGACTGACACCTTAATGCCAATAAAGTTGCTCATTTATGGACTGCT 80574

RESULT 14  
 AC123374/c  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-275L16, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 95 unordered pieces.  
 ACCESSION AC123374  
 VERSION AC123374.2 GI:21902551  
 KEYWORDS HTG, HTGS\_PHASE1.

SOURCE  
ORGANISM

Rattus norvegicus.  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE  
AUTHORS

1 (bases 1 to 182222)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
 Albrooks, S.L., Amaral, H.C., Are, C.R., Ayale, M., Banks, T.,  
 Barbata, J., Benton, J., Bimonte, K., Blomquist, K., Bonin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, F., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chaves, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dethorne, S.R., David, R.,  
 Delaney, K.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Devilla, M.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escott, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, J., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louissegh, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., M. Rindale, A., Martinez, E.,  
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newson, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokwenso, S., Ogih, M., Okunolu, G.,  
 Oranuyne, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickett, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolie, M., Ruiz, S., Savary, G.,  
 Scherz, E., Scott, G., Shen, H., Shooshari, N., Sisson, J.,  
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamer, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villa, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

TITLE  
JOURNAL

Direct Submission  
 Unpublished  
 Worley, K.C.

REFERENCE  
AUTHORS

2 (bases 1 to 182222)  
 Direct Submission  
 Submitted (29-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

TITLE  
JOURNAL

3 (bases 1 to 182222)  
 Direct Submission  
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS

On Jul 18, 2002 this sequence version replaced gi:21240277.

TITLE  
JOURNAL

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 Project Information  
 Center project name: GWLE  
 Center clone name: CH230-275L16  
 Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap; version 1.990129

## COMMENT

Consensus quality: 117047 bases at least Q40  
Consensus quality: 124820 bases at least Q30  
Consensus quality: 129800 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html))

- \* NOTE: This is a 'working draft' sequence. It currently consists of 95 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1061: contig of 1061 bp in length
2	1161: gap of unknown length
3	2208: contig of 1047 bp in length
4	2308: gap of unknown length
5	3792: contig of 1484 bp in length
6	3892: gap of unknown length
7	3993: contig of 1104 bp in length
8	4996: gap of unknown length
9	5097: contig of 1101 bp in length
10	6198: gap of unknown length
11	7376: contig of 1079 bp in length
12	7476: gap of unknown length
13	7477: contig of 1249 bp in length
14	8726: gap of unknown length
15	8825: gap of unknown length
16	10188: contig of 1363 bp in length
17	10288: gap of unknown length
18	11454: contig of 1166 bp in length
19	11455: gap of unknown length
20	11555: contig of 1619 bp in length
21	13173: gap of unknown length
22	13174: contig of 1159 bp in length
23	14432: gap of unknown length
24	14433: contig of 1266 bp in length
25	15798: gap of unknown length
26	15799: contig of 1097 bp in length
27	15899: gap of unknown length
28	16979: contig of 1081 bp in length
29	17079: gap of unknown length
30	17080: contig of 1568 bp in length
31	18647: gap of unknown length
32	18648: contig of 1097 bp in length
33	18748: gap of unknown length
34	19844: gap of unknown length
35	19845: contig of 1374 bp in length
36	21319: gap of unknown length
37	21418: contig of 1483 bp in length
38	22801: gap of unknown length
39	23001: contig of 1187 bp in length
40	23002: gap of unknown length
41	24188: contig of 1292 bp in length
42	24189: gap of unknown length
43	24289: contig of 1463 bp in length
44	25580: gap of unknown length
45	25581: contig of 1106 bp in length
46	25681: gap of unknown length
47	27143: contig of 1213 bp in length
48	27144: gap of unknown length
49	27244: contig of 1399 bp in length
50	28349: gap of unknown length
51	28350: contig of 1109 bp in length
52	28449: gap of unknown length
53	28450: contig of 1213 bp in length
54	28450: gap of unknown length
55	29663: contig of 1399 bp in length
56	29762: gap of unknown length
57	31161: contig of 1399 bp in length
58	31261: gap of unknown length
59	31262: contig of 1109 bp in length
60	32370: gap of unknown length
61	32371: contig of 1637 bp in length
62	34108: gap of unknown length
63	34207: contig of 1187 bp in length
64	34208: gap of unknown length
65	35394: contig of 1402 bp in length
66	35395: gap of unknown length
67	36896: contig of 1402 bp in length
68	36897: gap of unknown length
69	39359: contig of 1363 bp in length
70	39360: gap of unknown length
71	38460: contig of 1233 bp in length
72	39692: gap of unknown length
73	39693: contig of 1201 bp in length
74	39793: gap of unknown length

*	40994	41093: gap of unknown length
*	41094	43161: contig of 2068 bp in length
*	43162	43261: gap of unknown length
*	43262	44559: contig of 1198 bp in length
*	44460	44559: gap of unknown length
*	44360	45827: contig of 1268 bp in length
*	45028	45927: gap of unknown length
*	45328	47186: contig of 1259 bp in length
*	47187	47286: gap of unknown length
*	47287	48971: contig of 1685 bp in length
*	48972	49071: gap of unknown length
*	49072	50341: contig of 1270 bp in length
*	50342	50441: gap of unknown length
*	50442	51855: contig of 1414 bp in length
*	51856	51955: gap of unknown length
*	51956	53089: contig of 1134 bp in length
*	53090	53189: gap of unknown length
*	53190	54954: contig of 1765 bp in length
*	54955	55054: gap of unknown length
*	55055	55491: contig of 1437 bp in length
*	55492	55591: gap of unknown length
*	55592	58292: contig of 1701 bp in length
*	58293	58392: gap of unknown length
*	58393	59009: contig of 1517 bp in length
*	59010	60009: gap of unknown length
*	60010	61437: contig of 1428 bp in length
*	61438	61537: gap of unknown length
*	61538	64032: contig of 2495 bp in length
*	64033	64132: gap of unknown length
*	64133	65704: contig of 1572 bp in length
*	65705	65804: gap of unknown length
*	65805	67677: contig of 1873 bp in length
*	67678	67777: gap of unknown length
*	67778	69446: contig of 1669 bp in length
*	69447	69546: gap of unknown length
*	69547	70900: contig of 1354 bp in length
*	70901	71000: gap of unknown length
*	71001	73231: contig of 2231 bp in length
*	73232	73331: gap of unknown length
*	73332	74362: contig of 1031 bp in length
*	73332	74462: gap of unknown length
*	74363	76534: contig of 2072 bp in length
*	74463	76634: gap of unknown length
*	76535	78509: contig of 1875 bp in length
*	76635	78609: gap of unknown length
*	78510	80611: contig of 2002 bp in length
*	78610	80711: gap of unknown length
*	80712	82409: contig of 1698 bp in length

Query Match 34.2%; Score 1235.8; DB 2; Length 182222;  
Best Local Similarity 81.8%; Pred. No. 8.4e-280;  
Matches 1246; Conservative 0; Mismatches 278; Indels 0; G

Qy	1160	CTGTGTCGTCACTGGGCTTTGGCTCTGCTCGGAGGCCCTCTTTGGGGCCCGGCTGGG	1219
Db	89879	CGTGGTTCGTCACTGGGCTTTGGCTCTGCTCGGAGGCCCTCTTTGGGGCCCGGCTGGG	89820
Qy	1220	CCAAGGGCTTAGGGCCCATCCACTGAGTAGGTCGCTTGAGGGGATATGAGCGACCCCT	1279
Db	89819	CCAAGGGCTTAGGGCCCATCCACTGAGTAGGTCGCTTGAGGGGATATGAGCGACCCCT	89760
Qy	1280	CAGCGACTGCCCTGCCCTGGAAGGGTCCAGAAATGGTTGGCCAAATGAGAAATGCTGCTGC	1339
Db	89759	CAGCGACTGCCCTGCCCTGGAAGGGTCCAGAAATGGTTGGCCAAATGAGAAATGCTGCTGC	89700
Qy	1340	TGTCAGGTGCAATGTCCCTAAATGGGCTTTCAGAAATCAGGTGCGCTTGCGCTGGGGCG	1399
Db	89699	TGTCAGGTGCAATGTCCCTAAATGGGCTTTCAGAAATCAGGTGCGCTTGCGCTGGGGCG	89640
Qy	1400	TATCCCTGAGGAGGGGCTATTGAGGTGCAGGTGGAGTGAACGGGGTCCACGCTGGGG	1459
Db	89639	TATCCCTGAGGAGGGGCTATTGAGGTGCAGGTGGAGTGAACGGGGTCCACGCTGGGG	89580

[illegible]



us-09-924-946-1.rge

Wed Apr 2 09:13:55 2003

QY	2110	AGGAGTGACTGTAGGCTGCTGGGACACTACCGGCATGACATTGATTGCCAGTGGGTGG	2169
Db	2254	AGGCGATCACCATTGGCTGTGGGACATGTACCGCCATGACATGCATGCCAGTGGGTG	2313
QY	2170	ATATCACAGATGTGGGCCCGGGAATTATCTTCCAGTGAATTGTGAACCCCCCACTATG	2229
Db	2314	ACATCATGAGCTGGCCCTTGAGACTACCTGTTCCAGTTGTTATTAAACCCCACTTCG	2373
QY	2230	AAGTGGCAGAGTCAGATTTCTCCACAATATGCTGCAGTGGCGTGCAGAGTATGATGGGC	2289
Db	2374	AGGTTGCAGAAATCCGATTACTCCACAACATCATGAAATGCAGGAGCCGCTATGACGGCC	2433
QY	2290	ACCGGGTCTGGGTGCACAACCTGCCACACAGGGAATTCATACCCAGCCAACTGCAGAACTCT	2349
Db	2434	ACCCGACTTGGATGTACAACCTGCCACATAGGTGGTTCCTTCACGCCGAGAGACGGAATAA	2493
QY	2350	CCCTGGACGAGGAACAGCGGTCTCAGGAACAACC	2382
Db	2494	AGTTTGAGCACTTCAGCGGGCTCTTTAAACAACC	2526

Search completed: April 1, 2003, 07:03:51  
Job time : 19875 secs